

STIC-Biotech/ChemLib

131954

From: Whiteman, Brian
Sent: Wednesday, September 08, 2004 12:54 PM
To: STIC-Biotech/ChemLib
Subject: seq search

09/729,264 Welcher et al., 11/28/00

Please perform an olimoger search againts SEQ ID NOs: 1, 3 and 5.

If possible limit search to at least 16 nucleotides or more.

Thank you,

Brian Whiteman
Remsen, 2D14
mail box 2C18
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
(571) 272-0764

CREE

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: 9/13
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence :# _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 22:59:39 ; Search time 3446.95 Seconds
(without alignments)
10179.449 Million cell updates/sec

Title: US-09-729-264-1
Perfect score: 1175
Sequence: 1 ctgtctgccatctgaataa.....gtaatacaactgtatag 1175

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55024014

Minimum DB seq length: 16

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST.*

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pin.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	614	52.3	1201	13 BX399881	BX399881 BX399881
2	498	42.4	1201	13 BX358190	BX358190 BX358190
3	453	38.6	917	13 BX370118	BX370118 BX370118
4	443	37.7	941	13 BX403420	BX403420 BX403420

5	328	27.9	970	13	BX370209	BX370209
6	267	22.7	725	12	BG740428	BG740428
7	226	19.2	1148	13	BX370210	BX370210
8	149	12.7	216	12	BG206666	BG206666
9	146	12.4	622	29	AG069679	AG069679
10	146	12.4	677	29	AG107877	AG107877
11	110	9.4	1201	13	BX358189	BX358189
12	100	8.5	185	14	N47851	N47851
13	100	8.5	234	14	N93995	N93995
14	76	6.5	525	28	AQ403719	AQ403719
15	63	5.4	976	14	CK230652	CK230652
16	51	4.3	457	9	AJ003343	AJ003343
17	51	4.3	554	14	CB128910	CB128910
18	45	3.8	983	14	CK230612	CK230612
19	34	2.9	527	10	BE032610	BE032610
20	33	2.8	685	29	AG142221	AG142221
21	27	2.3	565	14	CD344913	CD344913
22	25	2.1	635	28	AQ541776	AQ541776
23	24	2.0	68	29	AB082057	AB082057
24	24	2.0	680	9	AA941851	AA941851
25	23	2.0	219	13	BQ862485	BQ862485
26	23	2.0	291	28	AZ411779	AZ411779
27	23	2.0	316	13	BQ864023	BQ864023
28	23	2.0	631	10	BB663870	BB663870
29	23	2.0	664	28	BH020641	BH020641
30	23	2.0	773	14	CB648901	CB648901
31	23	2.0	784	12	BG863468	BG863468
32	23	2.0	794	10	BF784177	BF784177
33	23	2.0	1510	11	AK086973	AK086973
34	22	1.9	167	29	CG475380	CG475380
35	22	1.9	171	29	CG480695	CG480695
36	22	1.9	188	13	BQ134412	BQ134412
37	22	1.9	196	29	CG480560	CG480560
38	22	1.9	219	29	CG479279	CG479279
39	22	1.9	223	29	CG615044	CG615044
40	22	1.9	232	29	CG481750	CG481750
41	22	1.9	234	29	CG481764	CG481764
42	22	1.9	238	29	CG566401	CG566401
43	22	1.9	287	10	BB343618	BB343618
44	22	1.9	328	29	CG476832	CG476832
45	22	1.9	340	14	Z46691	Z46691

ALIGNMENTS

RESULT 1
BX399881
LOCUS
DEFINITION
BX399881 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1085YA24 5-PRIME, mRNA sequence.
ACCESSION
BX399881
VERSION
BX399881.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 1201)
AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished (2001)
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10299.f, and
it belongs to a clone representative of this cluster. For more
information about this cluster and the virtual cDNA, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1085BA120Plkcluster=10299.f. Contact :
Peng Liang Email : fliang@lifetech.com URL :

BX399881 1201 bp mRNA linear EST 13-MAY-2003
BX399881 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1085YA24 5-PRIME, mRNA sequence.
BX399881
BX399881.1 GI:30622019
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10299.f, and
it belongs to a clone representative of this cluster. For more
information about this cluster and the virtual cDNA, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1085BA120Plkcluster=10299.f. Contact :
Peng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DI085BA12QPI.

FEATURES

source
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI085YA24"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 52.3%; Score 614; DB 13; Length 1201;
Best Local Similarity 100.0%; Pred. No. 1.3e-295;
Matches 614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 52 GTTCTGGGTCTGTAATGAGTCATAGAGGCCGCCAGAAATCAACAGTCCTGAAGGGCT 111
Db 272 GTTCTGGGTCTGTAATGAGTCATAGAGGCCGCCAGAAATCAACAGTCCTGAAGGGCT 331
QY 112 CCCAGGCTCGCTTCAACTGCACCGTCTCCAGGGCTGGAAGTCTCATCATGTGGGCTCTCA 171
Db 332 CCCAGGCTCGCTTCAACTGCACCGTCTCCAGGGCTGGAAGTCTCATCATGTGGGCTCTCA 391
QY 172 GTGACATGTGTGTGTTAAAGCTCAGGCCCATGAGGCCCATCATCATCAATGACCGGCTTCA 231
Db 392 GTGACATGTGTGTGTTAAAGCTCAGGCCCATGAGGCCCATCATCATCAATGACCGGCTTCA 451
QY 232 CCTCTCAGAGTACGACAGGGGGAACTTCACTCCGAGATGATCATCCCAATGTG 291
Db 452 CCTCTCAGAGTACGACAGGGGGAACTTCACTCCGAGATGATCATCCCAATGTG 511
QY 292 AGCCCATGATTCGGGGAAACATCAGATGACGCTCCAGAACAGTCGCTGCATGATCTG 351
Db 512 AGCCCATGATTCGGGGAAACATCAGATGACGCTCCAGAACAGTCGCTGCATGATCTG 571
QY 352 CTTACCTTACCGTCCAAAGTTATGGGAGAGTGTTTCAATCCAGTGTTAATCTTGTAGTCG 411
Db 572 CTTACCTTACCGTCCAAAGTTATGGGAGAGTGTTTCAATCCAGTGTTAATCTTGTAGTCG 631
QY 412 CTGAGATGACCTTGTGAAGTACTGTCTACCTCCACATGACCTGACCCGCTCCCGGATA 471
Db 632 CTGAGATGACCTTGTGAAGTACTGTCTACCTCCACATGACCTGACCCGCTCCCGGATA 691
QY 472 TTTCTCTGGAGCTCGGTCTCTCTGTCAGCCATTCAGCTATTATTTTGTTCGGAGCCCA 531
Db 692 TTTCTCTGGAGCTCGGTCTCTCTGTCAGCCATTCAGCTATTATTTTGTTCGGAGCCCA 751
QY 532 GCGACCTTCAAAGTGCAAGTGAAGTCTGCTGCTGACCCCAACAGCAATGGGACTTTGA 591
Db 752 GCGACCTTCAAAGTGCAAGTGAAGTCTGCTGCTGACCCCAACAGCAATGGGACTTTGA 811
QY 592 CTTGCGTGGTACTCGAGAGCTGAGCGCCGCAAGTCTGCAACTGTAAATCTCACTG 651
Db 812 CTTGCGTGGTACTCGAGAGCTGAGCGCCGCAAGTCTGCAACTGTAAATCTCACTG 871
QY 652 TGATTCGGTGTCCC 665
Db 872 TGATTCGGTGTCCC 885

RESULT 2

LOCUS BX358190 1201 bp mRNA linear EST 05-MAY-2003
DEFINITION BX358190 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI035YC01 5-PRIME, mRNA sequence.
ACCESSION BX358190
VERSION BX358190.1 GI:30372233
KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li.W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10299.f
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DI035BA01QPI.

FEATURES

source

1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI035YC01"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 42.4%; Score 498; DB 13; Length 1201;
Best Local Similarity 100.0%; Pred. No. 1.5e-237;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 678 GGTGTTTAAATATTCAGGTGTTATCAAGTTTACCGAGTTAGGTTTTCATTGCGCT 737
Db 64 GGTGTTTAAATATTCAGGTGTTATCAAGTTTACCGAGTTAGGTTTTCATTGCGCT 123
QY 738 ACTTGGGGCAAGTTTGGACTTGGACTAGCAGGCACCATCTCTGACGGCGACGTGACT 797
Db 124 ACTTGGGGCAAGTTTGGACTTGGACTAGCAGGCACCATCTCTGACGGCGACGTGACT 183
QY 798 CTTACAAATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 857
Db 184 CTTACAAATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
QY 858 TGTGTTTCTGCTGTGAGAGAAAGAGGATTTCTGTTTCAATTTCAAAGAAATCTGAA 917
Db 244 TGTGTTTCTGCTGTGAGAGAAAGAGGATTTCTGTTTCAATTTCAAAGAAATCTGAA 303
QY 918 AAAGAGAGACAAACAAAGAAACTGAGACAGAAAGTGGAAATGAAATCCGGCTACAAT 977
Db 304 AAAGAGAGACAAACAAAGAAACTGAGACAGAAAGTGGAAATGAAATCCGGCTACAAT 363
QY 978 TCAGATGAACAAAGAACCAACAGACACCGTCTCTCCCTCCCAAAATCTGTGAATCAGT 1037
Db 364 TCAGATGAACAAAGAACCAACAGACACCGTCTCTCCCTCCCAAAATCTGTGAATCAGT 423
QY 1038 GATCCTGMAACAAGAACACAGTAGCTGTGGCCCTCCCTCACCAGCGGGCTGATCAACGTCCA 1097
Db 424 GATCCTGMAACAAGAACACAGTAGCTGTGGCCCTCCCTCACCAGCGGGCTGATCAACGTCCA 483
QY 1098 CCCAGGCCAGCAAGTCAATCCACAGGCTTCTTTTAAATCTGGCCAGTCTGAGAAAGGTGAGT 1157
Db 484 CCCAGGCCAGCAAGTCAATCCACAGGCTTCTTTTAAATCTGGCCAGTCTGAGAAAGGTGAGT 543
QY 1158 AATCAACTGTAGTATAG 1175
Db 544 AATCAACTGTAGTATAG 561

RESULT 3


```

Db 526 AGCCGAGTATTCGGGGAAACATCAGATGCGCTCCAGAACAGTGCCTGCGATGATCTG 585
QY 352 CTTACTTACCGTCCAAAGTTATGGAGAGCTGTTCAATCCAGTGTAACTTGTAGTCG 411
Db 586 CTTACTTACCGTCCAAAGTTATGGAGAGCTGTTCAATCCAGTGTAACTTGTAGTCG 645
QY 412 CTGAGAAAGAACTTGTGAAGTTACTTGTCTACCTCCACACTGGACCGGCTCCCGGATA 471
Db 646 CTGAGAAAGAACTTGTGAAGTTACTTGTCTACCTCCACACTGGACCGGCTCCCGGATA 705
QY 472 TTTCTCGGAGCTCGGCTCTCCTG 494
Db 706 TTTCTCGGAGCTCGGCTCTCCTG 728

```

```

RESULT 5
BX370209 970 bp mRNA linear EST 08-MAY-2003
LOCUS
DEFINITION
BX370209 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1085YA24 5-PRIME, mRNA sequence.

```

```

ACCESSION
BX370209
VERSION
BX370209.1 GI:30451880
KEYWORDS
EST.
SOURCE
Homo sapiens (human)

```

ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE

```

1. (bases 1 to 970)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10299.f, and
it belongs to a clone representative of this cluster. For more
information about this cluster and the virtual cDNA, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAG039ZE05_CS03685_1&cluster=10299.f.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitrogenCorporation1600
Faraday Avenue Genoscope sequence ID : CS0BAG039ZE05_CS03685_1.

```

FEATURES

source

```

1. .970
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1085YA24"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

```

Query Match 27.9%; Score 328; DB 13; Length 970;
Best Local Similarity 100.0%; Pred. No. 1.7e-152;
Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 673 CTGGAGGTGGTAAATATTCAGGTGTTATTCAGTTTACCGAGTTTAGGTTTTCAT 732
Db 32 CTGGAGGTGGTAAATATTCAGGTGTTATTCAGTTTACCGAGTTTAGGTTTTCAT 91
QY 733 TGCCTACTTGGGCAAGTTGGACTTGGACTAGCAGGCACCATGCTTCTGACCCGACGT 792
Db 92 TGCCTACTTGGGCAAGTTGGACTTGGACTAGCAGGCACCATGCTTCTGACCCGACGT 151
QY 793 GTACTCTTCAATACCGTGTCTGCTGCGCGCTGTTGTTGGCTGCAACTGTGCT 852

```

```

Db 152 GTACTCTTCAATACCGTGTCTGCTGCGCGCTGTTGTTGGCTGCAACTGTGCT 211
QY 853 GCGGTTGTTGTTCTGCTGTAGAGAAAAGAGGATTTTCGTAATTCATATTTCAAAAGAAAT 912
Db 212 GCGGTTGTTGTTCTGCTGTAGAGAAAAGAGGATTTTCGTAATTCATATTTCAAAAGAAAT 271
QY 913 CTGAAAAGAGAGAACAAACAAAGAACTGAGACAGAAAGTGGAATGAAAATCTCGGCT 972
Db 272 CTGAAAAGAGAGAACAAACAAAGAACTGAGACAGAAAGTGGAATGAAAATCTCGGCT 331
QY 973 ACAATTCAGATCAACAAAAGACCACAGA 1000
Db 332 ACAATTCAGATCAACAAAAGACCACAGA 359

```

RESULT 6

```

BG740428 725 bp mRNA linear EST 15-MAY-2001
LOCUS
DEFINITION
602633817F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4778789 5',
mRNA sequence.

```

```

ACCESSION
BG740428
VERSION
BG740428.1 GI:14051081
KEYWORDS
EST.
SOURCE
Homo sapiens (human)

```

ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE

```

1. (bases 1 to 725)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10635 row: b column: 06
High quality sequence stop: 725.

```

FEATURES

Location/Qualifiers

```

1. .725
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4778789"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP_Skn3"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

```

ORIGIN

```

Query Match 22.7%; Score 267; DB 12; Length 725;
Best Local Similarity 100.0%; Pred. No. 5.3e-122;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 908 GAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGGAATGAAAATCTC 967
Db 379 GAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGGAATGAAAATCTC 438
QY 968 CGGCTACAAATTCAGATGAACAAAGAACACACACACCGCTTCTCTCCCTCCCAATCTCTG 1027
Db 439 CGGCTACAAATTCAGATGAACAAAGAACACACACACCGCTTCTCTCCCTCCCAATCTCTG 498
QY 1028 TGAATCCAGTGTCTTGAAACAAAGAACACAGTAGCTGTGGCCCTCTCTCACAGCGGCTGA 1087
Db 499 TGAATCCAGTGTCTTGAAACAAAGAACACAGTAGCTGTGGCCCTCTCTCACAGCGGCTGA 558
QY 1088 TCAAGTCCACCCAGCCAGCAAGTCATCCACAGGCTTCTTTAAATCGCCAGTCTCTGA 1147

```


Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 622)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimpesgsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. 622
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-060F12.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES
source
Query Match 12.4%; Score 146; DB 29; Length 622;
Best Local Similarity 100.0%; Pred.No. 1.9e-61;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 479 GGAGCTCGGTCTCTCGGTGAGCATTCAAGCTATTATTGTTCCGGAGCCCGAGCACT 538
Db 77 GGAGCTCGGTCTCTCGGTGAGCATTCAAGCTATTATTGTTCCGGAGCCCGAGCACT 136

QY 539 TCAAGTGCAGTGAGCATCTGGCTGACCCACAGAGCAATGGACTTTCACCTTGCCT 598
Db 137 TCAAGTGCAGTGAGCATCTGGCTGACCCACAGAGCAATGGACTTTCACCTTGCCT 196

QY 599 GGCTACCTGGAGAGCCTGAAGGCC 624
Db 197 GGCTACCTGGAGAGCCTGAAGGCC 222

RESULT 10
AG107877
LOCUS AG107877 677 bp DNA linear GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-112N07.F, genomic survey sequence.
ACCESSION AG107877
VERSION AG107877.1 GI:16728395
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 677)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimpesgsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,

Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. 677
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-112N07.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES
source
Query Match 12.4%; Score 146; DB 29; Length 677;
Best Local Similarity 100.0%; Pred.No. 1.9e-61;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 479 GGAGCTCGGTCTCTCGGTGAGCATTCAAGCTATTATTGTTCCGGAGCCCGAGCACT 538
Db 77 GGAGCTCGGTCTCTCGGTGAGCATTCAAGCTATTATTGTTCCGGAGCCCGAGCACT 136

QY 539 TCAAGTGCAGTGAGCATCTGGCTGACCCACAGAGCAATGGACTTTCACCTTGCCT 598
Db 137 TCAAGTGCAGTGAGCATCTGGCTGACCCACAGAGCAATGGACTTTCACCTTGCCT 196

QY 599 GGCTACCTGGAGAGCCTGAAGGCC 624
Db 197 GGCTACCTGGAGAGCCTGAAGGCC 222

RESULT 11
BX358189/c
LOCUS BX358189 1201 bp mRNA linear EST 05-MAY-2003
DEFINITION Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1035YC01 3-PRIME, mRNA sequence.
ACCESSION BX358189
VERSION BX358189.1 GI:30370198
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10299.f
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1035AB01NP1.
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1035YC01"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 9.4%; Score 110; DB 13; Length 1201;
Best Local Similarity 100.0%; Pred. No. 2.3e-43;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1066 GCCCTCCTCACAGCGGCTGATCAACGTCACCCAGCCAGCAAGTCATCCACAGGCTT 1125
|||||
Db 811 GCCCTCCTCACAGCGGCTGATCAACGTCACCCAGCCAGCAAGTCATCCACAGGCTT 752
|||||

QY 1126 CTTTAACTCGCCAGTCCTGAGAGGTCAGTAATACAACTGTAGTATAG 1175
|||||
Db 751 CTTTAACTCGCCAGTCCTGAGAGGTCAGTAATACAACTGTAGTATAG 702
|||||

RESULT 12

N47851 185 bp mRNA linear EST 14-FEB-1996
LOCUS YW95h05.r1 Soares_placenta_8to9weeks_2NbHP8to9W Homo sapiens cDNA
DEFINITION clone IMAGE:260025 5', mRNA sequence.
ACCESSION N47851
VERSION N47851.1 GI:1189017
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 185)
AUTHORS Hillier, L., Clark, N., Dubucque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: T7
High quality sequence stop: 150.

FEATURES

Location/Qualifiers

1..185
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3889731"
/db_xref="taxon:9606"
/clone="IMAGE:260025"
/dev_stage="two placentae: one from 8 weeks and another from 9 weeks post conception"
/lab_host="DHI0B (ampicillin resistant)"
/clone_lib="Soares_placenta_8to9weeks_2NbHP8to9W"
/notes="Organ: placenta; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo (dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGCGGCGGATTTTGTGTTT-3'], TGTTACCAATCTGAAGTGGAGCGCGGCGGATTTTGTGTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 8.5%; Score 100; DB 14; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.5e-38;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1076 CCAGCGGGCTGATCAACGTCACCCAGCCAGCAAGTCATCCACAGGCTTCTTTTAATCT 1135
|||||
Db 73 CCAGCGGGCTGATCAACGTCACCCAGCCAGCAAGTCATCCACAGGCTTCTTTTAATCT 132
|||||

QY 1136 GCCAGTCCTCGAGAGGTCAGTAATACAACTGTAGTATAG 1175
|||||
Db 133 GCCAGTCCTCGAGAGGTCAGTAATACAACTGTAGTATAG 172
|||||

RESULT 13

N93995 234 bp mRNA linear EST 05-APR-1996
LOCUS za66f09.r1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
DEFINITION IMAGE:297545 5', mRNA sequence.
ACCESSION N93995
VERSION N93995.1 GI:1266304
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 234)

AUTHORS Hillier, L., Clark, N., Dubucque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: mob.REGA+ET
High quality sequence stop: 159.

FEATURES Location/Qualifiers

1..234
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1242467"
/db_xref="taxon:9606"
/clone="IMAGE:297545"
/dev_stage="19 weeks"
/lab_host="DHI0B (ampicillin resistant)"
/clone_lib="Soares_fetal_lung_NbHL19W"
/notes="Organ: lung; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo (dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGCGGCGGATTTTGTGTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal heart library, Soares fetal heart NbHL19W."

ORIGIN

source

Query Match 8.5%; Score 100; DB 14; Length 234;
Best Local Similarity 100.0%; Pred. No. 1.6e-38;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1076 CCAGCGGGCTGATCAACGTCACCCAGCCAGCAAGTCATCCACAGGCTTCTTTTAATCT 1135
|||||
Db 113 CCAGCGGGCTGATCAACGTCACCCAGCCAGCAAGTCATCCACAGGCTTCTTTTAATCT 172
|||||

QY 1136 GGCCAGTCTCGAGAGGTCAGTAATACAACTGTAGTATAG 1175
|||||

Df	173	GCCAGTCTCGAAGGTCAGTCAATACACTGTAGTAGTAG	212
RESULT 14	AQ403719/c		
LOCUS	AQ403719	525 bp	linear
DEFINITION	HS_5049_A1_C11_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=625 Col=21 Row=E, genomic survey sequence.		
ACCESSION	AQ403719		
VERSION	AQ403719.1	GI:4414499	
KEYWORDS	GSS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,B.		
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)		
PUBLISHED	10449764		
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel.: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong [pieterejong.med.buffalo.edu]. Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu Plate: 625 row: E column: 21 Seq primer: T7 Class: BAC ends High quality sequence stop: 525. Location/Qualifiers 1..525 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /clone="Plate=625 Col=21 Row=E" /sex="male"		
FEATURES	source		
ORIGIN			
Query Match	6.5%;	Score 76;	DB 28; Length 525;
Best Local Similarity	100.0%;	Pred. No. 1.9e-26;	
Matches	76;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	1100	CAGGCCAGCAGTGATCATCCACAGGGTTCTTTTAATCTGCCAGTCTCTGAGAAGGTCAATAA	1159
Db	347	CAGGCCAGCAGTGATCATCCACAGGGTTCTTTTAATCTGCCAGTCTCTGAGAAGGTCAATAA	288
QY	1160	TACAACTGTAGTAGTAG	1175
Db	287	TACAACTGTAGTAGTAG	272
RESULT 15	CK230652		
LOCUS	CK230652	976 bp	mRNA linear
DEFINITION	ILLUMIGEN_MCO_1006 Katze_WMP12 Macaca mulatta cDNA 5' similar to		
human LOC150084 (Hs.422120), mRNA sequence.			
ACCESSION	CK230652		
VERSION	CK230652.1	GI:39636835	
KEYWORDS	EST.		
SOURCE	Macaca mulatta (rhesus monkey)		
ORGANISM	Macaca mulatta		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.		
AUTHORS	Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.		
TITLE	Large-scale Rhesus Macaque cDNA Sequencing Unpublished (2003)		
JOURNAL	Contact: C. Magness		
COMMENT	Illumigen Biosciences Inc. 2203 Airport Way S, Suite 450, Seattle, WA 98134, USA Tel: 2063780400 Fax: 2063780408 Email: cmagness@illumigen.com Sequenced on 2003.10.16. 695 Q20 bases. Assemblies in contig w/ 1 member(s). Contig contains 1 (0%) lib members. PCR Primers FORWARD: CCTCACTAAAGGAACAATAA BACKWARD: CACTATAGGGCGAATTGGGTA Insert Length: 976 Std Error: 0.00 Plate: CL000009 Row: C Column: 07 Seq primer: CCTCACTAAAGGAACAATAA POLYA=No.		
FEATURES	Location/Qualifiers		
1..976	/organism="Macaca mulatta" /mol_type="mRNA" /strain="Indian" /db_xref="taxon:9544" /sex="male" /dev_stage="newborn infant" /lab_host="E. coli SOLR" /clone_lib="Katze_WMP12" /note="Organ: placenta; Vector: Uni-ZAP XR; Site 1: EcORI I; Site 2: Xho I; Created from Stratagene ZAP-CDNA Synthesis kit (catalog #200400) and ZAP-CDNA Gigapack II Gold Cloning Kit (Catalog #200450)"		
ORIGIN			
Query Match	5.4%;	Score 63;	DB 14; Length 976;
Best Local Similarity	100.0%;	Pred. No. 7.2e-20;	
Matches	63;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	292	AGCCCAGTGTATTCGGGACATCAGATGCAGCTCCAGAACAGTGCCTCGATGATCTG	351
Db	581	AGCCCAGTGTATTCGGGACATCAGATGCAGCTCCAGAACAGTGCCTCGATGATCTG	
QY	352	CTT	354
Db	641	CTT	643
Search completed:	September 15, 2004,	06:43:43	
Job time :	3452.95 secs		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 14, 2004, 22:59:39 ; Search time 3426.42 Seconds
(without alignments)
10179.449 Million cell updates/sec

Title: US-09-729-264-3
Perfect score: 1168
Sequence: 1 agtgatcatggggcaggag.....gtaatacaactgtagtatag 1168

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0
Total number of hits satisfying chosen parameters: 55024014

Minimum DB seq length: 16
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pbg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	517	44.3	1201	13	BX399881
2	498	42.6	1201	13	BX358190
3	453	38.8	917	13	BX370118
4	361	30.9	941	13	BX403420

5	328	28.1	970	13	BX370209
6	267	22.9	725	12	BG740428
7	226	19.3	1148	13	BX370210
8	149	12.8	216	12	BG206666
9	146	12.5	622	29	AG069679
10	146	12.5	677	29	AG107877
c 11	110	9.4	1201	13	BX358189
c 12	100	8.6	185	14	N47851
c 13	100	8.6	234	14	N91995
c 14	76	6.5	525	28	AQ403719
15	63	5.4	976	14	CK230652
16	45	3.9	983	14	CK230612
17	34	2.9	527	10	BE032610
c 18	33	2.8	685	29	AG142221
c 19	27	2.3	565	14	CD344913
c 20	25	2.1	635	28	AQ541776
c 21	24	2.1	68	29	AB082057
c 22	24	2.1	680	9	AA941851
c 23	24	2.0	219	13	BQ862485
c 24	23	2.0	291	28	AZ411779
c 25	23	2.0	316	13	BQ864023
c 26	23	2.0	631	10	B8663870
c 27	23	2.0	664	28	BH020641
28	23	2.0	773	14	CB648901
29	23	2.0	784	12	BG863468
30	23	2.0	794	10	BF784177
31	23	2.0	1510	11	AK086973
32	22	1.9	167	29	CG475380
33	22	1.9	171	29	CG480895
34	22	1.9	188	13	BQ134412
35	22	1.9	196	29	CG480560
36	22	1.9	219	29	CG479279
37	22	1.9	223	29	CG615044
38	22	1.9	232	29	CG481750
39	22	1.9	234	29	CG481764
40	22	1.9	238	29	CG566401
c 41	22	1.9	287	10	BB343618
c 42	22	1.9	328	29	CG476832
c 43	22	1.9	340	14	Z46691
c 44	22	1.9	376	28	CC179595
c 45	22	1.9	424	9	AV409989

ALIGNMENTS

RESULT 1
BX399881
LOCUS
DEFINITION
BX399881 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CS0DI085YA24 5-PRIME, mRNA sequence.
ACCESSION
BX399881
VERSION
BX399881.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1201 bp mRNA linear EST 13-MAY-2003
BX399881 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CS0DI085YA24 5-PRIME, mRNA sequence.

Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10299.f, and
it belongs to a clone representative of this cluster. For more
information about this cluster and the virtual cDNA, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DI085BA12QPI&cluster=10299.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DI085BA12QP1.

FEATURES

source
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI085YA24"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 44.3%; Score 517; DB 13; Length 1201;
Best Local Similarity 99.8%; Pred. No. 2e-245;
Matches 567; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 91 AGTCCTGAAGGGCTCCAGGCTCGCTTCAACTGCACCGCTCCAGGGCTGGAAGCTCAT 150
Db 318 AGTCCTGAAGGGCTCCAGGCTCGCTTCAACTGCACCGCTCCAGGGCTGGAAGCTCAT 377
QY 151 CATGTGGGCTCTCAAGTACATGATGCTGCTAAAGCGTCAAGGCCCATGAGGCCCATCATCAC 210
Db 378 CATGTGGGCTCTCAAGTACATGATGCTGCTAAAGCGTCAAGGCCCATGAGGCCCATCATCAC 437
QY 211 CAATGACCGCTTCACTCTCAGAGGTACACAGGCGGGAACCTTCACTCGGAGATGAT 270
Db 438 CAATGACCGCTTCACTCTCAGAGGTACACAGGCGGGAACCTTCACTCGGAGATGAT 497
QY 271 CATCCACAATGTGGAGCCAGTGTATTCGGGGAACATCAGATGCAGCTCCAGAACAGTCG 330
Db 498 CATCCACAATGTGGAGCCAGTGTATTCGGGGAACATCAGATGCAGCTCCAGAACAGTCG 557
QY 331 CTGATGATGATCTGCTTACCTTACCGTCCAAGTATATGGAGAGCTGTTCATTTCCCAAGTGT 390
Db 558 CTGATGATGATCTGCTTACCTTACCGTCCAAGTATATGGAGAGCTGTTCATTTCCCAAGTGT 617
QY 391 TAATCTTGTAGTGTGAGAAATGAACCTTGTGAAGTACTTGTACCCCTCACTGGAC 450
Db 618 TAATCTTGTAGTGTGAGAAATGAACCTTGTGAAGTACTTGTACCCCTCACTGGAC 677
QY 451 CTGGCTCCCGGATATTTCTGGAGCTCGGTCTCTCTGGTCAAGCTTCAAGCTATTATTT 510
Db 678 CTGGCTCCCGGATATTTCTGGAGCTCGGTCTCTCTGGTCAAGCTTCAAGCTATTATTT 737
QY 511 TGTTCGGAGCCGAGCGACTTAAAGTCAGTGAGCATCTCTGGCTCTGACCCCAAGAG 570
Db 738 TGTTCGGAGCCGAGCGACTTAAAGTCAGTGAGCATCTCTGGCTCTGACCCCAAGAG 797
QY 571 CAATGGGACTTTCAGTTCGTGCTACCTTGAAGAGCTTGAAGCCCGGAGCCGAGTCTGCAAC 630
Db 798 CAATGGGACTTTCAGTTCGTGCTACCTTGAAGAGCTTGAAGCCCGGAGCCGAGTCTGCAAC 857
QY 631 TGTAAATCTCACTGTGATTCGGTGTCCC 658
Db 858 TGTAAATCTCACTGTGATTCGGTGTCCC 885

RESULT 2

BX358190
LOCUS BX358190 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0DI035YC01 5-PRIME, mRNA sequence.
ACCESSION BX358190
VERSION BX358190.1 GI:30372233
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)

COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10299.f
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DI035BA01QP1.

FEATURES

source

1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI035YC01"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 42.6%; Score 498; DB 13; Length 1201;
Best Local Similarity 100.0%; Pred. No. 5.6e-236;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 671 GGTGGTATTATTAATTCAGGCTGATTAATCAAGTTTACCGAGTTTAGTTTTTCATTGCGCT 730
Db 64 GGTGGTATTATTAATTCAGGCTGATTAATCAAGTTTACCGAGTTTAGTTTTTCATTGCGCT 123
QY 731 ACTTGGGGCAAAAGTTGGACTTGGACTAGCAGGCACCATGCTTCTGACGCCGACGTGACT 790
Db 124 ACTTGGGGCAAAAGTTGGACTTGGACTAGCAGGCACCATGCTTCTGACGCCGACGTGACT 183
QY 791 CTTACAATACGCTGCTGCTGCTGCCGCGCTGCTTGTGTGGCTGCACACTGCTGCTGCCGT 850
Db 184 CTTACAATACGCTGCTGCTGCTGCCGCGCTGCTTGTGTGGCTGCACACTGCTGCTGCCGT 243
QY 851 TGTGTGTTCTGCTGTAGAGAAAAGAGGATTTGCTATTCAATTTCAAAAGAAATCTGAA 910
Db 244 TGTGTGTTCTGCTGTAGAGAAAAGAGGATTTGCTATTCAATTTCAAAAGAAATCTGAA 303
QY 911 AAAGAGAAGACAAACAAAGAACTGAGACAGAAAGTGGAAATGAAATCCCGGCTACAAT 970
Db 304 AAAGAGAAGACAAACAAAGAACTGAGACAGAAAGTGGAAATGAAATCCCGGCTACAAT 363
QY 971 TCAGATGAACAAAGACACACAGCCGCTTCTCCCTCCCAATCTCTGTAATCCAGT 1030
Db 364 TCAGATGAACAAAGACACACAGCCGCTTCTCCCTCCCAATCTCTGTAATCCAGT 423
QY 1031 GATCCTGAACAAAGAAACAGTAGCTGTGGCCCTCTCCACAGCGGCTCATCAAGTCCA 1090
Db 424 GATCCTGAACAAAGAAACAGTAGCTGTGGCCCTCTCCACAGCGGCTCATCAAGTCCA 483
QY 1091 CCCAGGCGAGCAAGTCATCCACAGGCTCTTTTAAATCTGGCCAGTCTTGAGAGGTCAGT 1150
Db 484 CCCAGGCGAGCAAGTCATCCACAGGCTCTTTTAAATCTGGCCAGTCTTGAGAGGTCAGT 543
QY 1151 AATACAACCTGTAGTATAG 1168
Db 544 AATACAACCTGTAGTATAG 561

RESULT 3

BX370118

LOCUS

DEFINITION

BX370118 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI076YN22 5-PRIME, mRNA sequence.

```
ACCESSION BX370118
VERSION BX370118.1 GI:30447910
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2450.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAG0372E05_CS03493_2&cluster=2250.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BAG0372E05_CS03493_2.
FEATURES
source
Location/Qualifiers
1..917
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1076YN22"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 38.8%; Score 453; DB 13; Length 917;
Best Local Similarity 99.8%; Pred. No. 1.3e-213;
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 665 ACTGGAGTGGTATTATATCCAGGTGTTATATCAAGTTACGAGTTAGGTTTTCA 724
Db 31 ACTGGAGTGGTATTATATCCAGGTGTTATATCAAGTTACGAGTTAGGTTTTCA 90
QY 725 TTGCCTACTTGGGCAAGTTGACTTGGACTAGCAGSCACCTGCTTCTGACGCCGACG 784
Db 91 TTGCCTACTTGGGCAAGTTGACTTGGACTAGCAGSCACCTGCTTCTGACGCCGACG 150
QY 785 TGTAATCTTACAATACGCTGCTGCTGCGCGCGTGTGTTGTTGCTGCACTGCTGC 844
Db 151 TGTAATCTTACAATACGCTGCTGCTGCGCGCGTGTGTTGTTGCTGCACTGCTGC 210
QY 845 TCGCGTGTGTTTCTGCTAGAGAAAAGAGGATTTGTAATCAATTTCAAAAGAAA 904
Db 211 TCGCGTGTGTTTCTGCTAGAGAAAAGAGGATTTGTAATCAATTTCAAAAGAAA 270
QY 905 TCTGAAAAGAGAGAGACAAAGAACTGAGACAGAAAGTGAATGAAACTCCGGC 964
Db 271 TCTGAAAAGAGAGAGACAAAGAACTGAGACAGAAAGTGAATGAAACTCCGGC 330
QY 965 TACAATTCAGATGAACAAAGACCAAGACACACGCTTCTCTCCCTCCCAATCTGTGA 1024
Db 331 TACAATTCAGATGAACAAAGACCAAGACACACGCTTCTCTCCCTCCCAATCTGTGA 390
QY 1025 TCCAGTGATCTGAACAAAGAAACAGTAGCTGTGGCCCTCTCCACAGCGGGTGATCAA 1084
Db 391 TCCAGTGATCTGAACAAAGAAACAGTAGCTGTGGCCCTCTCCACAGCGGGTGATCAA 450
QY 1085 CGTCCACCCAGGCCAGCAGTATCCACAGGCTTCTTTTAACTGGCCAGTCTGAGAAG 1144
Db 451 CGTCCACCCAGGCCAGCAGTATCCACAGGCTTCTTTTAACTGGCCAGTCTGAGAAG 510
```

```
QY 1145 GTCAGTAATACAACTGTAGTATAG 1168
|||||
Db 511 GTCAGTAATACAACTGTAGTATAG 534
|||||

RESULT 4
LOCUS BX403420
DEFINITION BX403420 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1085YA24 5-PRIME, mRNA sequence.
ACCESSION BX403420
VERSION BX403420.1 GI:30607302
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
10299.f, and it belongs to a clone representative of this cluster.
For more information about this cluster and the virtual cDNA, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS5AA017ZE02RM1&cluster=10299.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS5AA017ZE02RM1.
FEATURES
source
Location/Qualifiers
1..941
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1085YA24"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 30.9%; Score 361; DB 13; Length 941;
Best Local Similarity 100.0%; Pred. No. 7.1e-168;
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 91 AGTCCTGAAGGGCTCCAGGCTCCACTGCACCGTCTCCAGGGCTGGAAGCTCAT 150
Db 332 AGTCCTGAAGGGCTCCAGGCTCCACTGCACCGTCTCCAGGGCTGGAAGCTCAT 391
QY 151 CATGTGGGCTCTCAGTGACATGTGTGTGTAAAGCTGAGGCCCATGAGCCCATCATC 210
Db 392 CATGTGGGCTCTCAGTGACATGTGTGTGTAAAGCTGAGGCCCATGAGCCCATCATC 451
QY 211 CAATGACCGCTTCCACTCTCAGAGGTACGACAGGGGGAACTTCACTTCGAGATGAT 270
Db 452 CAATGACCGCTTCCACTCTCAGAGGTACGACAGGGGGAACTTCACTTCGAGATGAT 511
QY 271 CATCCCAATGTGGAGCCCAAGTATTCGGGGAAACATCAGATGCAGCTCCAGAACAGTCG 330
Db 512 CATCCCAATGTGGAGCCCAAGTATTCGGGGAAACATCAGATGCAGCTCCAGAACAGTCG 571
QY 331 CTTGCAATGATCTGCTTACCTTACCGTCCAAAGTTATGGAGAGCTCTTCACTCCAGTGT 390
Db 572 CTTGCAATGATCTGCTTACCTTACCGTCCAAAGTTATGGAGAGCTCTTCACTCCAGTGT 631
QY 391 TAATCTTGTAGTCGCTGAGAAATCAACCTTGTGAGTACTTGTCTACCTTCACATGGAC 450
```

Db 632 TAATCTGTAGTCGTGAGTAACCTTGTGAAGTACTTGTCTACCCCTCACATGGAC 691
 QY 451 C 451
 Db 692 C 692

RESULT 5
 BX370209
 LOCUS
 DEFINITION BX370209 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CS0DI085YA24 5-PRIME, mRNA sequence.
 ACCESSION BX370209
 VERSION
 KEYWORDS
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 970)
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 JOURNAL
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91060 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 10299.f, and
 it belongs to a clone representative of this cluster. For more
 information about this cluster and the virtual cDNA, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0BAG039ZE05_CS03685_1&cluster=10299.f.
 Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS0BAG039ZE05_CS03685_1.

FEATURES
 source
 1..970
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DI085YA24"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 28.1%; Score 328; DB 13; Length 970;
 Best Local Similarity 100.0%; Pred. No. 1.8e-151;
 Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 666 CTGAGGTGGTATTAAATATCCAGGTGATTATCAAGTTTACCAGTTTAGTTTTCAT 725
 Db 32 CTGAGGTGGTATTAAATATCCAGGTGATTATCAAGTTTACCAGTTTAGTTTTCAT 91
 QY 726 TGCCTACTTGGGGCAAGTTGGACTTGGACTAGCAGGCACCATGCTTCTGACGCCGAGT 785
 Db 92 TGCCTACTTGGGGCAAGTTGGACTTGGACTAGCAGGCACCATGCTTCTGACGCCGAGT 151
 QY 786 GTACTCTTAATAGCTGCTGCTGCTGCCCGCTGCTGTTGGTGGCTGCACTGCTGCT 845
 Db 152 GTACTCTTAATAGCTGCTGCTGCTGCCCGCTGCTGTTGGTGGCTGCACTGCTGCT 211
 QY 846 GCCGTGTTGTTCTGCTGTAGACAAAAAGAGGATTTTCGATTCAATTTCAAAGAAGAT 905
 Db 212 GCCGTGTTGTTCTGCTGTAGACAAAAAGAGGATTTTCGATTCAATTTCAAAGAAGAT 271
 QY 906 CTGAAAAAGAGAAGCAAAACAAAGAACTGAGACAGAAAGTGGAAATGAAAATCCGGCT 965

Db 272 CTGAAAAAGAGAAGCAAAACAAAGAACTGAGACAGAAAGTGGAAATGAAAATCCGGCT 331
 QY 966 ACAATTCAGATGAACAAAAAGACCACAGA 993
 Db 332 ACAATTCAGATGAACAAAAAGACCACAGA 359

RESULT 6
 BG740428
 LOCUS
 DEFINITION BG740428 725 bp mRNA linear EST 15-MAY-2001
 mRNA sequence.
 ACCESSION BG740428
 VERSION
 KEYWORDS
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 725)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: James Cleaver, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
 Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10635 row: b column: 06
 High quality sequence stop: 725.

FEATURES
 source
 1..725
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4778789"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI CGAP Skn3"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.5kb. Library constructed by Life
 Technologies. Note: this is a NCI CGAP Library."

ORIGIN
 Query Match 22.9%; Score 267; DB 12; Length 725;
 Best Local Similarity 100.0%; Pred. No. 3.7e-121;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 901 GAAATCTGAAAAAGAGAGAACAAAGAACTGAGACAGAAAGTGGAAATGAAAACTC 960
 Db 379 GAAATCTGAAAAAGAGAGAACAAAGAACTGAGACAGAAAGTGGAAATGAAAACTC 438
 QY 961 CGGCTACAATTCAGATGAACAAAAAGACCACAGACACCGCTTCTCTCCCTCCCAATCTG 1020
 Db 439 CGGCTACAATTCAGATGAACAAAAAGACCACAGACACCGCTTCTCTCCCTCCCAATCTG 498
 QY 1021 TGAATCCAGTATCTCTGAACAAAGAAACAGTAGCTGGCCCTCTCACCAGCGGGCTGA 1080
 Db 499 TGAATCCAGTATCTCTGAACAAAGAAACAGTAGCTGGCCCTCTCACCAGCGGGCTGA 558
 QY 1081 TCAAGTCCACCCAGGCGAGCAAGTCATCCACAGGCTTCTTTTAACTGCGCAGTCTCTGA 1140
 Db 559 TCAAGTCCACCCAGGCGAGCAAGTCATCCACAGGCTTCTTTTAACTGCGCAGTCTCTGA 618
 QY 1141 GAAAGTCAGTAATCAACTGTAGTATA 1167
 Db 619 GAAAGTCAGTAATCAACTGTAGTATA 645

```

RESULT 7
BX370210      1148 bp      mRNA      linear      EST 08-MAY-2003
LOCUS
DEFINITION    clone CS0DI085YA24 5-PRIME, mRNA sequence.
ACCESSION    BX370210
VERSION      BX370210.1 GI:30451881
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10299.f, and
it belongs to a clone representative of this cluster. For more
information about this cluster and the virtual cDNA, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAG039ZE05_CS03685_2&cluster=10299.f.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BAG039ZE05_CS03685_2.
FEATURES
source
Location/Qualifiers
1..1148
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI085YA24"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match      19.3%; Score 226; DB 13; Length 1148;
Best Local Similarity 100.0%; Pred. No. 9.9e-101;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 658 CCAAGACACTGGAGGTGGTATTATATTCAGGTGTATTATCAAGTTACCGAGTTTAGG 717
Db 14 CCAAGACACTGGAGGTGGTATTATATTCAGGTGTATTATCAAGTTTACCGAGTTTAGG 73

QY 718 TTTTTCATTGCTACTTTGGGCAAAAGTTGGACTTGGACTAGCAGCCACCATGCTTCTGAC 777
Db 74 TTTTTCATTGCTACTTTGGGCAAAAGTTGGACTTGGACTAGCAGCCACCATGCTTCTGAC 133

QY 778 GCGACGTGACTCTTCAATACAGCTGCTGCTGCGCGCGTGTGTTGTTGTGCTGCA 837
Db 134 GCGACGTGACTCTTCAATACAGCTGCTGCTGCGCGCGTGTGTTGTTGTGCTGCA 193

QY 838 CTGCTGCTGCGTGTGTTGTTCTGCTGTAGAGAAAAGAGATT 883
Db 194 CTGCTGCTGCGTGTGTTGTTCTGCTGTAGAGAAAAGAGATT 239

RESULT 8
BG206666
LOCUS
DEFINITION    R5T6117 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION    BG206666
VERSION      BG206666.1 GI:13728353
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10299.f, and
it belongs to a clone representative of this cluster. For more
information about this cluster and the virtual cDNA, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAG039ZE05_CS03685_2&cluster=10299.f.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BAG039ZE05_CS03685_2.
FEATURES
source
Location/Qualifiers
1..1148
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI085YA24"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match      19.3%; Score 226; DB 13; Length 1148;
Best Local Similarity 100.0%; Pred. No. 9.9e-101;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 658 CCAAGACACTGGAGGTGGTATTATATTCAGGTGTATTATCAAGTTACCGAGTTTAGG 717
Db 14 CCAAGACACTGGAGGTGGTATTATATTCAGGTGTATTATCAAGTTTACCGAGTTTAGG 73

QY 718 TTTTTCATTGCTACTTTGGGCAAAAGTTGGACTTGGACTAGCAGCCACCATGCTTCTGAC 777
Db 74 TTTTTCATTGCTACTTTGGGCAAAAGTTGGACTTGGACTAGCAGCCACCATGCTTCTGAC 133

QY 778 GCGACGTGACTCTTCAATACAGCTGCTGCTGCGCGCGTGTGTTGTTGTGCTGCA 837
Db 134 GCGACGTGACTCTTCAATACAGCTGCTGCTGCGCGCGTGTGTTGTTGTGCTGCA 193

QY 838 CTGCTGCTGCGTGTGTTGTTCTGCTGTAGAGAAAAGAGATT 883
Db 194 CTGCTGCTGCGTGTGTTGTTCTGCTGTAGAGAAAAGAGATT 239

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 216)
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
Offenbacher,J., Danzig,J. and Ducar,M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
11329013
COMMENT
Contact: Scott J. Cain
Athersys, Inc. Ave, Cleveland, OH 44115, USA
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 166.
FEATURES
source
Location/Qualifiers
1..216
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

ORIGIN
Query Match      12.8%; Score 149; DB 12; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.3e-62;
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1011 CCAATCTGTGAATCCAGTGATCTCTGAACAAAGAACAGTAGTGTGGCCCTCTCACC 1070
Db 21 CCAATCTGTGAATCCAGTGATCTCTGAACAAAGAACAGTAGTGTGGCCCTCTCACC 80

QY 1071 AGCGGGCTGATCACGTCACCCAGGCCAGCAAGTCTCCACAGGCTCTTTTATCTGG 1130
Db 81 AGCGGGCTGATCACGTCACCCAGGCCAGCAAGTCTCCACAGGCTCTTTTATCTGG 140

QY 1131 CCAAGTCTGAGAAGGTGAGTAATACAACT 1159
Db 141 CCAAGTCTGAGAAGGTGAGTAATACAACT 169

RESULT 9
AG069679
LOCUS
DEFINITION    Pan troglodytes DNA, clone: PTB-060F12.F, genomic survey sequence.
ACCESSION    AG069679
VERSION      AG069679.1 GI:16621481
KEYWORDS
SOURCE
ORGANISM      Pan troglodytes (chimpanzee)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished
2 (bases 1 to 622)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

```

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: chimbes@gscc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)

COMMENT
 Clones are derived from the chimpanzee BAC library PTB This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.

PRIMERS

Sequencing: -21M13

LIBRARY

Vector : pKS145

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

1. .622

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/clone="PTB-060F12.F"

/sex="male"

/cell_type="lymphoblast"

/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

Query Match

Best Local Similarity 12.5%; Score 146; DB 29; Length 622;

Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 472 GGAGCTCGGTCTCTGTCAGCCATTCAAGCTATTATTGTTCCGGAGCCGACCT 531

Db 77 GGAGCTCGGTCTCTGTCAGCCATTCAAGCTATTATTGTTCCGGAGCCGACCT 136

QY 532 TCAAGTGCATGAGCATCTGCTCTGACCCACAGCAATGGGACTTTGACTTGCCT 591

Db 137 TCAAGTGCATGAGCATCTGCTCTGACCCACAGCAATGGGACTTTGACTTGCCT 196

QY 592 GGTACTCTGAAGAGCTGAAGGCC 617

Db 197 GGTACTCTGAAGAGCTGAAGGCC 222

RESULT 10

AG107877

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Reference

Authors

Title

Journal

Comment

Primers

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

1. .677

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/clone="PTB-112N07.F"

/sex="male"

/cell_type="lymphoblast"

/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

Query Match

Best Local Similarity 12.5%; Score 146; DB 29; Length 677;

Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 472 GGAGCTCGGTCTCTGTCAGCCATTCAAGCTATTATTGTTCCGGAGCCGACCT 531

Db 77 GGAGCTCGGTCTCTGTCAGCCATTCAAGCTATTATTGTTCCGGAGCCGACCT 136

QY 532 TCAAGTGCATGAGCATCTGCTCTGACCCACAGCAATGGGACTTTGACTTGCCT 591

Db 137 TCAAGTGCATGAGCATCTGCTCTGACCCACAGCAATGGGACTTTGACTTGCCT 196

QY 592 GGTACTCTGAAGAGCTGAAGGCC 617

Db 197 GGTACTCTGAAGAGCTGAAGGCC 222

RESULT 11

BX358189/c

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Comment

Primers

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

1. .677

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/clone="PTB-112N07.F"

/sex="male"

/cell_type="lymphoblast"

/clone_lib="PTB Chimpanzee Male BAC Library"

Query Match

Best Local Similarity 12.5%; Score 146; DB 29; Length 677;

Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 472 GGAGCTCGGTCTCTGTCAGCCATTCAAGCTATTATTGTTCCGGAGCCGACCT 531

Db 77 GGAGCTCGGTCTCTGTCAGCCATTCAAGCTATTATTGTTCCGGAGCCGACCT 136

QY 532 TCAAGTGCATGAGCATCTGCTCTGACCCACAGCAATGGGACTTTGACTTGCCT 591

Db 137 TCAAGTGCATGAGCATCTGCTCTGACCCACAGCAATGGGACTTTGACTTGCCT 196

QY 592 GGTACTCTGAAGAGCTGAAGGCC 617

Db 197 GGTACTCTGAAGAGCTGAAGGCC 222

RESULT 11

BX358189/c

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Comment

Primers

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

Sequencing: -21M13

Library

Vector

```

QY 1059 GCCCTCTCACAGCGGGGTGATCAACGTCACCCAGGCCAGCAAGTCATCCACAGGCTT 1118
Db 811 GCCCTCTCACAGCGGGGTGATCAACGTCACCCAGGCCAGCAAGTCATCCACAGGCTT 752
QY 1119 CTTTAAATCTGCCAGTCCTGAGGAGGTCAGTAATACAACTAGTATATAG 1168
Db 751 CTTTAAATCTGCCAGTCCTGAGGAGGTCAGTAATACAACTAGTATATAG 702

RESULT 12
LOCUS N47851 185 bp mRNA linear EST 14-FEB-1996
DEFINITION yw95h05.r1 Soares_placenta 8to9weeks 2NDHP8to9W Homo sapiens cDNA
clone IMAGE:260025 5', mRNA sequence.
ACCESSION N47851
VERSION N47851.1 GI:1189017
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 185)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: T7
High quality sequence stop: 150.
FEATURES
source
1..185
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3889731"
/db_xref="taxon:9606"
/clone="IMAGE:260025"
/dev_stage="Two placentae; one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_placenta 8to9weeks 2NDHP8to9W"
/note="Organ: placenta; Vector: pT7u3D (Pharmacia) with a
modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'-
TGTTACCAATCTGAAGTGGAGCGCGCGATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."
ORIGIN
Query Match 8.6%; Score 100; DB 14; Length 185;
Best Local Similarity 100.0%; Pred. No. 3e-38;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1069 CCAGCGGGCTGATCAAGTCCACCGCCAGCAAGTCATCCACAGGCTTCTTTTAACT 1128
Db 73 CCAGCGGGCTGATCAAGTCCACCGCCAGCAAGTCATCCACAGGCTTCTTTTAACT 132
QY 1129 GGCCAGTCCTGAGAAGGTGAGTAATACAACTAGTATATAG 1168
Db 133 GGCCAGTCCTGAGAAGGTGAGTAATACAACTAGTATATAG 172

Query Match 8.6%; Score 100; DB 14; Length 185;
Best Local Similarity 100.0%; Pred. No. 3e-38;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1069 CCAGCGGGCTGATCAAGTCCACCGCCAGCAAGTCATCCACAGGCTTCTTTTAACT 1128
Db 113 CCAGCGGGCTGATCAAGTCCACCGCCAGCAAGTCATCCACAGGCTTCTTTTAACT 172
QY 1129 GGCCAGTCCTGAGAAGGTGAGTAATACAACTAGTATATAG 1168
Db 173 GGCCAGTCCTGAGAAGGTGAGTAATACAACTAGTATATAG 212

Query Match 8.6%; Score 100; DB 14; Length 234;
Best Local Similarity 100.0%; Pred. No. 3.2e-38;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1069 CCAGCGGGCTGATCAAGTCCACCGCCAGCAAGTCATCCACAGGCTTCTTTTAACT 1128
Db 113 CCAGCGGGCTGATCAAGTCCACCGCCAGCAAGTCATCCACAGGCTTCTTTTAACT 172
QY 1129 GGCCAGTCCTGAGAAGGTGAGTAATACAACTAGTATATAG 1168
Db 173 GGCCAGTCCTGAGAAGGTGAGTAATACAACTAGTATATAG 212

RESULT 14
LOCUS AQ403719 525 bp DNA linear GSS 13-MAR-1999
DEFINITION HS 5049 A1.C11.T7A RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=625 Col=21 Row=E, genomic survey sequence.

```

```

RESULT 13
LOCUS N93995 234 bp mRNA linear EST 05-APR-1996
DEFINITION zaf6f09.r1 Soares_fetal_lung NbHL19W Homo sapiens cDNA clone
IMAGE:297545 5', mRNA sequence.
ACCESSION N93995
VERSION N93995.1 GI:1266304
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 234)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: mob.REGA+ET
High quality sequence stop: 159.
FEATURES
source
1..234
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1242467"
/db_xref="taxon:9606"
/clone="IMAGE:297545"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_fetal_lung NbHL19W"
/note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCAATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. This library was constructed
from the same fetus as the fetal heart library. Soares
fetal heart NbHL19W."
ORIGIN
Query Match 8.6%; Score 100; DB 14; Length 234;
Best Local Similarity 100.0%; Pred. No. 3.2e-38;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1069 CCAGCGGGCTGATCAAGTCCACCGCCAGCAAGTCATCCACAGGCTTCTTTTAACT 1128
Db 113 CCAGCGGGCTGATCAAGTCCACCGCCAGCAAGTCATCCACAGGCTTCTTTTAACT 172
QY 1129 GGCCAGTCCTGAGAAGGTGAGTAATACAACTAGTATATAG 1168
Db 173 GGCCAGTCCTGAGAAGGTGAGTAATACAACTAGTATATAG 212

RESULT 14
LOCUS AQ403719 525 bp DNA linear GSS 13-MAR-1999
DEFINITION HS 5049 A1.C11.T7A RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=625 Col=21 Row=E, genomic survey sequence.

```

ACCESSION AQ403719
 VERSION AQ403719.1 GI:4414499
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 525)
 AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 PUBMED 10449764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: <http://www.htsc.washington.edu>
 Plate: 625 row: E column: 21
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 525.

FEATURES
 source
 1..525
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="plate=625 Col=21 Row=E"
 /sex="male"
 /clone_lib="RPCI-11 Human Male BAC Library"
 /notes="vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at EcoRI sites"

ORIGIN
 Query Match 6.5%; Score 76; DB 28; Length 525;
 Best Local Similarity 100.0%; Pred. No. 3.2e-26;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1093 CAGCCAGCAGTCATCCAGGCTTCTTTAATCTGCCAGTCTCTCAGAGGTCAGTAA 1152
 Db |||||
 347 CAGCCAGCAGTCATCCAGGCTTCTTTAATCTGCCAGTCTCTCAGAGGTCAGTAA 288
 |||||
 QY 1153 TACAACTGTAGTATAG 1168
 Db |||||
 287 TACAACTGTAGTATAG 272
 |||||

RESULT 15
 CK230652 976 bp mRNA linear EST 09-DEC-2003
 ILLUMIGEN MCQ 1006 Katze_MNPL2 Macaca mulatta cDNA 5' similar to
 human LOC150084 (Hs.422120), mRNA sequence.
 ACCESSION CK230652
 VERSION CK230652.1 GI:39636835
 KEYWORDS EST.
 SOURCE Macaca mulatta (rhesus monkey)
 ORGANISM Macaca mulatta
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Cercopitheciinae; Macaca.
 1 (bases 1 to 976)
 Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
 Large-scale Rhesus Macaque cDNA Sequencing
 Unpublished (2003)
 Contact: C. Magness
 Illumigen Biosciences Inc.
 2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
 Tel: 2063780400
 Fax: 2063780408
 Email: cmagness@illumigen.com
 Sequenced on 2003.10.16. 695 Q20 bases. Assemblies in contig w/ 1 member(s). Contig contains 1 (0%) lib members.
 PCR Primers
 FORWARD: CCTCACTAAAGGGAACAAA
 BACKWARD: CACTATAGGCGAATTGGTA
 Insert Length: 976 Std Error: 0.00
 Plate: CL000009 row: C column: 07
 Seq primer: CCTCACTAAAGGGAACAAA
 POLYA=NO.

FEATURES
 source
 1..976
 Location/Qualifiers
 /organism="Macaca mulatta"
 /mol_type="mRNA"
 /strain="Indian"
 /db_xref="taxon:9544"
 /sex="male"
 /dev_stage="newborn infant"
 /lab_host="E. coli SOLR"
 /clone_lib="Katze_MNPL2"
 /notes="Organ: placenta; Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: Xho I; Created from Stratagene ZAP-CDNA Synthesis kit (catalog #200400) and ZAP-CDNA Gigapack III Gold Cloning Kit (Catalog #200450)"

ORIGIN
 Query Match 5.4%; Score 63; DB 14; Length 976;
 Best Local Similarity 100.0%; Pred. No. 1.1e-19;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 285 AGCCCACTGATTCGGGAACATCAGATGCAGCCTCCAGAACAGTCGCTCATGATCTG 344
 Db |||||
 581 AGCCCACTGATTCGGGAACATCAGATGCAGCCTCCAGAACAGTCGCTCATGATCTG 640
 |||||
 QY 345 CTT 347
 Db |||||
 641 CTT 643
 |||||

Search completed: September 15, 2004, 06:43:44
 Job time : 3427.42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 22:59:39 ; Search time 3637.63 Seconds
(without alignments)
10179.449 Million cell updates/sec

Title: US-09-729-264-5

Perfect score: 1240

Sequence: 1 aggtgtgagtcgacgaaca.....gtaatacaactgttagtag 1240

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55024014

Minimum DB seq length: 16

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_esti:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pug:*
27: em_gss_vri:*
28: gb_ges1:*
29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	614	49.5	1201	13 BX399881	BX399881 BX399881
2	504	40.6	917	13 BX370118	BX370118 BX370118
3	447	36.0	1201	13 BX358190	BX358190 BX358190
4	443	35.7	941	13 BX403420	BX403420 BX403420

5	336	27.1	970	13 BX370209	BX370209 BX370209
6	226	18.2	1148	13 BX370210	BX370210 BX370210
7	216	17.4	725	12 BG740428	BG740428 BG740428
8	149	12.0	216	12 BG206666	BG206666 RST26117
9	146	11.8	622	29 AG069679	AG069679 Pan trogl
10	146	11.8	677	29 AG107877	AG107877 Pan trogl
11	110	8.9	1201	13 BX358189	BX358189 BX358189
12	100	8.1	185	14 N47851	N47851 YW95H05.r1
13	100	8.1	234	14 N93995	N93995 za66f09.r1
14	80	6.5	1201	13 BX396896	BX396896 BX396896
15	76	6.1	525	28 AQ403719	AQ403719 HS_5049_A
16	63	5.1	976	14 CK230652	CK230652 ILJUMIGEN
17	45	3.6	983	14 CK230612	CK230612 ILJUMIGEN
18	34	2.7	527	10 BE032610	BE032610 132035_MA
19	33	2.7	685	29 AG142221	AG142221 Pan trogl
20	32	2.6	1201	13 BX396897	BX396897 BX396897
21	27	2.2	565	14 CD344913	CD344913 EtBSTee75
22	25	2.0	228	29 CE729707	CE729707 tigr-988-
23	25	2.0	635	28 AO541776	AQ541776 RCI-11-3
24	25	2.0	834	29 CC588638	CC588638 CH240_387
25	24	1.9	68	29 AB082057	AB082057 Drosophila
26	24	1.9	680	9 AA941851	AA941851 LD27161.5
27	23	1.9	219	13 BQ862485	BQ862485 QGC21D13
28	23	1.9	291	28 AZ411779	AZ411779 IM0185B04
29	23	1.9	316	13 BQ864023	BQ864023 QSC25108
30	23	1.9	631	10 BB663870	BB663870 BB663870
31	23	1.9	664	28 BH020641	BH020641 LB612a.d
32	23	1.9	721	28 AQ316148	AQ316148 RCI11-10
33	23	1.9	773	14 CB648901	CB648901 OSJNB121
34	23	1.9	784	12 BG863468	BG863468 602796941
35	23	1.9	794	10 BF784177	BF784177 602108039
36	23	1.9	824	28 BZ842183	BZ842183 CH240_213
37	23	1.9	1510	11 AK086973	AK086973 Mus muscu
38	22	1.8	167	29 CG475380	CG475380 OST12736
39	22	1.8	171	29 CG480695	CG480695 OST12736
40	22	1.8	186	13 BQ134412	BQ134412 1091016H0
41	22	1.8	198	29 CG480560	CG480560 OST12523
42	22	1.8	219	29 CG479279	CG479279 OST9980_M
43	22	1.8	223	29 CG615044	CG615044 OST303854
44	22	1.8	232	29 CG481750	CG481750 OST14380
45	22	1.8	234	29 CG481764	CG481764 OST14399

ALIGNMENTS

RESULT 1
BX399881
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BX399881 1201 bp mRNA linear EST 13-MAY-2003
BX399881 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1085YA24 5-PRIME, mRNA sequence.

BX399881
BX399881
EST
Homo sapiens (human)

BX399881
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1201)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 10299.f, and

it belongs to a clone representative of this cluster. For more

information about this cluster and the virtual cDNA, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0D1085BA120P1&cluster=10299.f. Contact :

Feng Liang Email: fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DI085BA12QPL.

FEATURES

source

```

1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon.9606"
/clone="CS001085YA24"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

```
Query Match          49.5%; Score 614; DB 13; Length 1201;
Best Local Similarity 100.0%; Pred. NO. 1.3e-294;
Watchdog 614 Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

117	GTCTGGGTCTGGTAATGAAGTCATAGAAGGCCCCAGAAATGCAACAGTCTCTGAAGGCT	176	GTCTGGGTCTGGTAATGAAGTCATAGAAGGCCCCAGAAATGCAACAGTCTCTGAAGGCT
272	GTCTGGGTCTGGTAATGAAGTCATAGAAGGCCCCAGATGCAACAGTCTCTGAAGGCT	331	GTCTGGGTCTGGTAATGAAGTCATAGAAGGCCCCAGATGCAACAGTCTCTGAAGGCT
177	CCAGGCTCGCTTCAACTGCACCGCTCTCCAGGGGCTGGAAGCTCATATGTGGGCTCTCA	236	CCAGGCTCGCTTCAACTGCACCGCTCTCCAGGGGCTGGAAGCTCATATGTGGGCTCTCA
332	CCAGGCTCGCTTCAACTGCACCGCTCTCCAGGGGCTGGAAGCTCATATGTGGGCTCTCA	391	CCAGGCTCGCTTCAACTGCACCGCTCTCCAGGGGCTGGAAGCTCATATGTGGGCTCTCA
237	GTGACATGTFGTGCTTAAGCGTTCAGGCCCATGAGGCCCATCATCAACCATGACCGCTTCA	296	GTGACATGTFGTGCTTAAGCGTTCAGGCCCATGAGGCCCATCATCAACCATGACCGCTTCA
392	GTGACATGTFGTGCTTAAGCGTTCAGGCCCATGAGGCCCATCATCAACCATGACCGCTTCA	451	GTGACATGTFGTGCTTAAGCGTTCAGGCCCATGAGGCCCATCATCAACCATGACCGCTTCA
297	CTCTCAGAGGTACGACACGAGGCGGGAACTTCACTTCGAGATGATCATCCACAATGTGG	356	CTCTCAGAGGTACGACACGAGGCGGGAACTTCACTTCGAGATGATCATCCACAATGTGG
452	CTCTCAGAGGTACGACACGAGGCGGGAACTTCACTTCGAGATGATCATCCACAATGTGG	511	CTCTCAGAGGTACGACACGAGGCGGGAACTTCACTTCGAGATGATCATCCACAATGTGG
357	AGCCCACTGATTTCGGGGAAACATCAGATGCAGCCTCCAGAACAGTGCCTGCATGCATCTG	416	AGCCCACTGATTTCGGGGAAACATCAGATGCAGCCTCCAGAACAGTGCCTGCATGCATCTG
512	AGCCCACTGATTTCGGGGAAACATCAGATGCAGCCTCCAGAACAGTGCCTGCATGCATCTG	571	AGCCCACTGATTTCGGGGAAACATCAGATGCAGCCTCCAGAACAGTGCCTGCATGCATCTG
417	CTTACCTTACCGTCCAAGTTATGGGAGAGCTGTTCAATCCCAAGTGTTAATCTTTAGTCTG	476	CTTACCTTACCGTCCAAGTTATGGGAGAGCTGTTCAATCCCAAGTGTTAATCTTTAGTCTG
572	CTTACCTTACCGTCCAAGTTATGGGAGAGCTGTTCAATCCCAAGTGTTAATCTTTAGTCTG	631	CTTACCTTACCGTCCAAGTTATGGGAGAGCTGTTCAATCCCAAGTGTTAATCTTTAGTCTG
477	CTCAGAAATGAACCTTGTGAAGTTACTTTGTCTACCTTCACACTGGACCCCGGCTCCCGGATA	536	CTCAGAAATGAACCTTGTGAAGTTACTTTGTCTACCTTCACACTGGACCCCGGCTCCCGGATA
632	CTCAGAAATGAACCTTGTGAAGTTACTTTGTCTACCTTCACACTGGACCCCGGCTCCCGGATA	691	CTCAGAAATGAACCTTGTGAAGTTACTTTGTCTACCTTCACACTGGACCCCGGCTCCCGGATA
537	TTTCTCTGGGAGCTCGGCTCTCCTGGTCAGCCATTCAAGCTATATTTTGTTCGGAGCCCCA	596	TTTCTCTGGGAGCTCGGCTCTCCTGGTCAGCCATTCAAGCTATATTTTGTTCGGAGCCCCA
692	TTTCTCTGGGAGCTCGGCTCTCCTGGTCAGCCATTCAAGCTATATTTTGTTCGGAGCCCCA	751	TTTCTCTGGGAGCTCGGCTCTCCTGGTCAGCCATTCAAGCTATATTTTGTTCGGAGCCCCA
597	GGACACTTCAAAGTCAGATGAGCATCTGTGGCTCTGACCCACAGAGCAATGGGACTTTGA	656	GGACACTTCAAAGTCAGATGAGCATCTGTGGCTCTGACCCACAGAGCAATGGGACTTTGA
752	GGACACTTCAAAGTCAGATGAGCATCTGTGGCTCTGACCCACAGAGCAATGGGACTTTGA	811	GGACACTTCAAAGTCAGATGAGCATCTGTGGCTCTGACCCACAGAGCAATGGGACTTTGA
657	CTTGTGCTGGCTTACTCTGGAAGAGCTGGAAGCGCCGCAAGTCTGCAACTGTAAATCTCACTG	716	CTTGTGCTGGCTTACTCTGGAAGAGCTGGAAGCGCCGCAAGTCTGCAACTGTAAATCTCACTG
812	CTTGTGCTGGCTTACTCTGGAAGAGCTGGAAGCGCCGCAAGTCTGCAACTGTAAATCTCACTG	871	CTTGTGCTGGCTTACTCTGGAAGAGCTGGAAGCGCCGCAAGTCTGCAACTGTAAATCTCACTG
717	TGATTCGGTGTCCC	730	TGATTCGGTGTCCC
872	TGATTCGGTGTCCC	885	TGATTCGGTGTCCC

RESULT 2
BX370118
LOCUS
DEFINITION BX370118 Homo sapiens PLACENTA CP 25-NORMALIZED Homo sapiens CDNA
clone CSOD1076VN22 5-PRIME, mRNA sequence.
ACCSSION BX370118
VERSION BX370118.1 GI:30447910
KEYWORDS EST.

```

RESULT 3
BX358190
LOCUS
DEFINITION
BX358190 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI035YC01 5-PRIME, mRNA sequence.
ACCESSION
BX358190
VERSION
BX358190.1 GI:30372233
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10299.f
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DI035AB01QP1.
LOCATION/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI035YC01"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
FEATURES
source
Query Match 36.0%; Score 447; DB 13; Length 1201;
Best Local Similarity 99.8%; Pred. No. 2.6e-211;
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
ORIGIN
743 GGTGGTATTATATTCAGGTGATTATCAAGTTTACCGAGTTTAGTGTTCATTGCGCT 802
64 GGTGGTATTATATTCAGGTGATTATCAAGTTTACCGAGTTTAGTGTTCATTGCGCT 123
803 ACTTGGGGCAAGTTGGACTTGGACTAGCAGGACCAATGCTTGTGACCGCAGCTGTACT 862
124 ACTTGGGGCAAGTTGGACTTGGACTAGCAGGACCAATGCTTGTGACCGCAGCTGTACT 183
863 CTTTACATAGCTGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 922
184 CTTTACATAGCTGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
923 TGTGTTTCTGCTGTAGAGAAAGAGGATTTTCGTAATTCATTAATTCATTAATTCATTAAT 982
244 TGTGTTTCTGCTGTAGAGAAAGAGGATTTTCGTAATTCATTAATTCATTAATTCATTAAT 1042
983 AAAGAGAGCAACAAAGAACTGAGACAGAAAGTGGAAATGAAATTCGAGTCAAT 1042
304 AAAGAGAGCAACAAAGAACTGAGACAGAAAGTGGAAATGAAATTCGAGTCAAT 363
1043 TCAGATGAACAAAGAACCAAGAACTGAGACAGAAAGTGGAAATGAAATTCGAGTCAAT 1102
364 TCAGATGAACAAAGAACCAAGAACTGAGACAGAAAGTGGAAATGAAATTCGAGTCAAT 423
1103 GATCTCTGAACAAAGAACCAAGTGTGCGCTTCTCACCAGCGGGCTGATCAAGTCCA 1162
424 GATCTCTGAACAAAGAACCAAGTGTGCGCTTCTCACCAGCGGGCTGATCAAGTCCA 483

```

```

RESULT 4
BX403420
LOCUS
DEFINITION
BX403420 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI085YA24 5-PRIME, mRNA sequence.
ACCESSION
BX403420
VERSION
BX403420.1 GI:30607302
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 941)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
10299.f, and it belongs to a clone representative of this cluster.
For more information about this cluster and the virtual cDNA, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS5AA017ZE02RM1&cluster=10299.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS5AA017ZE02RM1.
LOCATION/Qualifiers
1. 941
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI085YA24"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
FEATURES
source
Query Match 35.7%; Score 443; DB 13; Length 941;
Best Local Similarity 100.0%; Pred. No. 2.5e-209;
Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ORIGIN
117 GTTCTGGGTCTGTTAAGTCAAGTCAAGAGGCGCCCATGAGCCCATCATCAATGACCGCTTCA 176
286 GTTCTGGGTCTGTTAAGTCAAGTCAAGAGGCGCCCATGAGCCCATCATCAATGACCGCTTCA 345
177 CCAGGCTCGTCTCAATCGACCGTCTCCAGGGCTGGAGCTCATCATGTGGGCTCTCA 236
346 CCAGGCTCGTCTCAATCGACCGTCTCCAGGGCTGGAGCTCATCATGTGGGCTCTCA 405
237 GTGACATGTTGTTGTTAAGTCAAGTCAAGGCGCCCATGAGCCCATCATCAATGACCGCTTCA 296
406 GTGACATGTTGTTGTTAAGTCAAGTCAAGGCGCCCATGAGCCCATCATCAATGACCGCTTCA 465
297 CCTCTCAGAGGTACGACGACGCGGGAACTTCACCTCGGAGATGATCATCAATGTGG 356
466 CCTCTCAGAGGTACGACGACGCGGGAACTTCACCTCGGAGATGATCATCAATGTGG 525
357 AGCCCGAGTGATTCGGGGAAACATCATGATGAGCGCTTCAGAACAGTGCCTGCATGATCTG 416

```

Db	526	AGCCAGTGATTCGGGGAAACATCAGATGAGCCTCAGAACAGTCGCCCTCATGATGATCTG	585
Qy	417	CTTACTTACCGTCCAGGTTATGGAGAGCTGTTCAITCCCAGTGTAAATCTTGTAGTCG	476
Db	586	CTTACTTACCGTCCAGGTTATGGAGAGCTGTTCAITCCCAGTGTAAATCTTGTAGTCG	645
Qy	477	CTGAGAATGAACCTTGTGAAGTTACTGTCTACCTCACCCTGACATGGACCCGGCTCCCGGATA	536
Db	646	CTGAGAATGAACCTTGTGAAGTTACTGTCTACCTCACCCTGACATGGACCCGGCTCCCGGATA	705
Qy	537	TTTCTCGGAGCTCGGTCCTCG	559
Db	706	TTTCTCGGAGCTCGGTCCTCG	728
RESULT 5			
BX370209			
LOCUS			
DEFINITION			
BX370209 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA			
clone CS0D1085YA24 5-PRIME, mRNA sequence.			
ACCESSION			
BX370209			
VERSION			
BX370209.1 GI:30451880			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
REFERENCE			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 970)			
Li, W. B., Gruber, C., Jessee, J., and Polayes, D.			
Full-length cDNA libraries and normalization			
Unpublished (2001)			
Contact: Genoscope			
Genoscope - Centre National de Sequencage			
BP 191 91006 EVRY cedex - France			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
Library was constructed by Life Technologies, a division of			
Invitrogen. This sequence belongs to sequence cluster 10299.f, and			
it belongs to a clone representative of this cluster. For more			
information about this cluster and the virtual cDNA, see			
http://www.genoscope.cns.fr/			
cgi-bin/cluster.cgi?seq=CS0BAG039ZE05_CS03685_1&cluster=10299.f.			
Contact : Feng Liang Email : fliang@lifetech.com URL :			
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600			
Faraday Avenue Genoscope sequence ID : CS0BAG039ZE05_CS03685_1.			
FEATURES			
Location/Qualifiers			
1..970			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CS0D1085YA24"			
/tissue type="PLACENTA COT 25-NORMALIZED"			
/note="1st strand cDNA was primed with a NotI-oligo (dr)			
primer. Five prime end enriched, double-strand cDNA was			
digested with Not I and cloned into the Not I and EcoR V			
sites of the pCMVSPORT 6 vector. Library was normalized."			
ORIGIN			
Query Match			
Best Local Similarity 27.1%; Score 336; DB 13; Length 970;			
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	738	CTGGAGGTGGTATTATATCCAGGTGATTTACAGTTTACCGAGTTAGTTTTCAT	797
Db	32	CTGGAGGTGGTATTATATCCAGGTGATTTACAGTTTACCGAGTTAGTTTTCAT	91
Qy	798	TGCCTACTTGGGCAAGTTGGACTTGGACTAGCAGGACCATGCTTCTGACCGCAGCT	857
Db	92	TGCCTACTTGGGCAAGTTGGACTTGGACTAGCAGGACCATGCTTCTGACCGCAGCT	151
Qy	858	GTACTCTTACAATACGCTGCTGCTGCCCGCTGTTGTTGTGCTGCAACTCGTGTCT	917

Db	152	GTACTCTTACAATACGCTGCTGCTGCCCGCTGTTGTTGTGGCTGCAACTGCTGCT	211
Qy	918	GCCGTTGTTGTTTCTGTAGAGAAAAGAGGATTTTCGATTCAATTTCAAAAAGAAAT	977
Db	212	GCCGTTGTTGTTTCTGTAGAGAAAAGAGGATTTTCGATTCAATTTCAAAAAGAAAT	271
Qy	978	CTGAAAAAGAGAAGCAAAACAAAGAACTGAGACAGAAAGTGAATGAAACTCCGGCT	1037
Db	272	CTGAAAAAGAGAAGCAAAACAAAGAACTGAGACAGAAAGTGAATGAAACTCCGGCT	331
Qy	1038	ACAATTCAGATCAACAAAAGACCACAGAAACCGCTT	1073
Db	332	ACAATTCAGATCAACAAAAGACCACAGAAACCGCTT	367
RESULT 6			
BX370210			
LOCUS			
DEFINITION			
BX370210 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA			
clone CS0D1085YA24 5-PRIME, mRNA sequence.			
ACCESSION			
BX370210			
VERSION			
BX370210.1 GI:30451881			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
REFERENCE			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 1148)			
Li, W. B., Gruber, C., Jessee, J., and Polayes, D.			
Full-length cDNA libraries and normalization			
Unpublished (2001)			
Contact: Genoscope			
Genoscope - Centre National de Sequencage			
BP 191 91006 EVRY cedex - France			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
Library was constructed by Life Technologies, a division of			
Invitrogen. This sequence belongs to sequence cluster 10299.f, and			
it belongs to a clone representative of this cluster. For more			
information about this cluster and the virtual cDNA, see			
http://www.genoscope.cns.fr/			
cgi-bin/cluster.cgi?seq=CS0BAG039ZE05_CS03685_2&cluster=10299.f.			
Contact : Feng Liang Email : fliang@lifetech.com URL :			
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600			
Faraday Avenue Genoscope sequence ID : CS0BAG039ZE05_CS03685_2.			
FEATURES			
Location/Qualifiers			
1..1148			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CS0D1085YA24"			
/tissue type="PLACENTA COT 25-NORMALIZED"			
/note="1st strand cDNA was primed with a NotI-oligo (dr)			
primer. Five prime end enriched, double-strand cDNA was			
digested with Not I and cloned into the Not I and EcoR V			
sites of the pCMVSPORT 6 vector. Library was normalized."			
ORIGIN			
Query Match			
Best Local Similarity 18.2%; Score 226; DB 13; Length 1148;			
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	730	CCAAGACACTGGAGTGGTATTAATATTCAGAGTGTATATCAAGTTTACCGAGTTTAGG	789
Db	14	CCAAGACACTGGAGTGGTATTAATATTCAGAGTGTATTAATCAAGTTTACCGAGTTTAGG	73
Qy	790	TTTTTCATTCGCTACTTCGGGCAAGTTGGACTTCGACTAGCAGGACCATGCTTCTGAC	849
Db	74	TTTTTCATTCGCTACTTCGGGCAAGTTGGACTTCGACTAGCAGGACCATGCTTCTGAC	133
Qy	850	GCCGACGTGTACTCTTACAATAACGCTGCTGCTGCCCGCTGTTGTTGTGGCTGCAA	909
Db	134	GCCGACGTGTACTCTTACAATAACGCTGCTGCTGCCCGCTGTTGTTGTGGCTGCAA	193

```

QY      910 CTGCTGCTGCCGTTGTTGTTCTGCTGTAGAGAAAAAGAGGATTT 955
Db      194 CTGCTGCTGCCGTTGTTGTTCTGCTGTAGAGAAAAAGAGGATTT 239

RESULT 7
LOCUS   BG740428
DEFINITION 602633817F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4778789 5',
mrna sequence.
ACCESSION BG740428
VERSION   EST.
KEYWORDS  BG740428.1 GI:14051081
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 725)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs@mail.nih.gov
          Tissue Procurement: James Cleaver, M.D.
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
          Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: L14M10635 row: b column: 06
          High quality sequence stop: 725.
          Location/Qualifiers
            1..725
              /organism="Homo sapiens"
              /mol_type="mrna"
              /db_xref="taxon:9606"
              /clone="IMAGE:4778789"
              /lab_host="DH10B (T1 phage-resistant)"
              /clone_lib="NCI CGAP Skn3"
              /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
              Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
              Average insert size 1.5kb. Library constructed by Life
              Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 17.4%; Score 216; DB 12; Length 725;
Best Local Similarity 99.6%; Pred. No. 4.1e-96;
Matches 266; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      973 GAAATCTGAAAAAGAGAAGACAAACAAAGAACTGTAGACAGAAAGTGAATGAAACTC 1032
Db      379 GAAATCTGAAAAAGAGAAGACAAACAAAGAACTGTAGACAGAAAGTGAATGAAACTC 438

QY      1033 CGSCTCAATTCAGATGACAAAGAACCCAGAAAACCGCTTCTCTCCCTCCCAAACTCTG 1092
Db      439 CGSCTCAATTCAGATGACAAAGAACCCAGAACCGCTTCTCTCCCTCCCAAACTCTG 498

QY      1093 TGAATCCAGTGTATCTTGAACAAAGAAACAGTAGCTGTGGCCCTCTCCACAGCGGCTGA 1152
Db      499 TGAATCCAGTGTATCTTGAACAAAGAAACAGTAGCTGTGGCCCTCTCCACAGCGGCTGA 558

QY      1153 TCAAGTCCACCCAGGCCAGCAGTAGTCATCCACAGGCTTCTTTTAATCTGGCCAGTCTCTGA 1212
Db      559 TCAAGTCCACCCAGGCCAGCAGTAGTCATCCACAGGCTTCTTTTAATCTGGCCAGTCTCTGA 618

QY      1213 GAAGGTCAGTAATACAACTGTAGTATA 1239
Db      619 GAAGGTCAGTAATACAACTGTAGTATA 645

RESULT 8
LOCUS   AG069679
DEFINITION Pan troglodytes DNA, clone: PTB-060F12.F, genomic survey sequence.
ACCESSION AG069679
VERSION   AG069679.1 GI:16621481
KEYWORDS  GSS.
SOURCE    Pan troglodytes (chimpanzee)
ORGANISM  Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1
AUTHORS   Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

```

```

BG206666
LOCUS    BG206666
DEFINITION RST26117 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG206666
VERSION   BG206666.1 GI:13728353
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 216)
AUTHORS   Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
          Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,
          Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,
          Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K.,
          Offenbacher, J., Danzig, J. and Ducar, M.
          Creation of genome-wide protein expression libraries using random
          activation of gene expression
          Nat. Biotechnol. 19 (5), 440-445 (2001)
JOURNAL   21227151
MEDLINE   11329013
PUBMED    11329013
COMMENT   Contact: Scott J. Cain
          Athersys, Inc.
          3201 Carnegie Ave, Cleveland, OH 44115, USA
          Tel: 216 431 9900
          Fax: 216 361 9596
          Email: scai@atersys.com
          High quality sequence stop: 166.
          Location/Qualifiers
            1..216
              /organism="Homo sapiens"
              /mol_type="mrna"
              /db_xref="taxon:9606"
              /cell_line="HT1080"
              /clone_lib="Athersys RAGE Library"
              /note="See 'Creation of Genome-wide Protein Expression
              Libraries using Random Activation of Gene Expression',
              Nature Biotechnology, in press. Note that even though the
              cell type indicated is HT1080, since a random activation
              method was used, these sequence tags are not necessarily
              expressed in HT1080 under normal circumstances."

ORIGIN
Query Match 12.0%; Score 149; DB 12; Length 216;
Best Local Similarity 100.0%; Pred. NO. 8.1e-63;
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1083 CCAATCTCTGTGAATCCAGTGATCCTCAACAAAGAAACAGTAGTGTGGCCCTCTCACC 1142
Db      21 CCAATCTCTGTGAATCCAGTGATCCTCAACAAAGAAACAGTAGTGTGGCCCTCTCACC 80

QY      1143 AGCGGCTGATCAACGTCCACCCAGCCAGCAAGTCATCCACAGGCTTCTTTTAATCTGG 1202
Db      81 AGCGGCTGATCAACGTCCACCCAGCCAGCAAGTCATCCACAGGCTTCTTTTAATCTGG 140

QY      1203 CCAGTCTCTGAGAAGTCAAGTATACAACT 1231
Db      141 CCAGTCTCTGAGAAGTCAAGTATACAACT 169

RESULT 9
LOCUS    AG069679
DEFINITION Pan troglodytes DNA, clone: PTB-060F12.F, genomic survey sequence.
ACCESSION AG069679
VERSION   AG069679.1 GI:16621481
KEYWORDS  GSS.
SOURCE    Pan troglodytes (chimpanzee)
ORGANISM  Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1
AUTHORS   Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

```

Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 622)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/;
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. .622
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-060F12.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match 11.8%; Score 146; DB 29; Length 622;
Best Local Similarity 100.0%; Pred. No. 3.3e-61;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 544 GGAGCTCGGTCCTCGTCAGCCATTCAAGCTATTATTGTTCCGGAGCCCGACCT 603
|||||
DB 77 GGAGCTCGGTCCTCGTCAGCCATTCAAGCTATTATTGTTCCGGAGCCCGACCT 136
QY 604 TCAAGTGCAGTGCAGTCTCTGCTGAGCCATTCAAGCTATTATTGTTCCGGAGCCCGACCT 663
|||||
DB 137 TCAAGTGCAGTGCAGTCTCTGCTGAGCCATTCAAGCTATTATTGTTCCGGAGCCCGACCT 196
QY 664 GGCTACCTGGAGAGCTCTGAAGGCC 689
|||||
DB 197 GGCTACCTGGAGAGCTCTGAAGGCC 222
RESULT 10
AG107877 677 bp DNA linear GSS 03-NOV-2001
LOCUS Pan troglodytes DNA, clone: PTB-112N07.F, genomic survey sequence.
DEFINITION AG107877
ACCESSION AG107877
VERSION AG107877.1 GI:16728395
KEYWORDS Pan troglodytes (chimpanzee)
SOURCE Pan troglodytes
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 677)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/;
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. .677
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-112N07.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match 11.8%; Score 146; DB 29; Length 677;
Best Local Similarity 100.0%; Pred. No. 3.3e-61;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 544 GGAGCTCGGTCCTCGTCAGCCATTCAAGCTATTATTGTTCCGGAGCCCGACCT 603
|||||
DB 77 GGAGCTCGGTCCTCGTCAGCCATTCAAGCTATTATTGTTCCGGAGCCCGACCT 136
QY 604 TCAAGTGCAGTGCAGTCTCTGCTGAGCCATTCAAGCTATTATTGTTCCGGAGCCCGACCT 663
|||||
DB 137 TCAAGTGCAGTGCAGTCTCTGCTGAGCCATTCAAGCTATTATTGTTCCGGAGCCCGACCT 196
QY 664 GGCTACCTGGAGAGCTCTGAAGGCC 689
|||||
DB 197 GGCTACCTGGAGAGCTCTGAAGGCC 222
RESULT 11
BX358189/c 1201 bp mRNA linear EST 05-MAY-2003
LOCUS BX358189 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
DEFINITION clone CSOD1035YC01 3-PRIME, mRNA sequence.
ACCESSION BX358189
VERSION BX358189.1 GI:30370198
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seq@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10299.f
Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CSOD1035AB01NPI.
1. .1201
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1035YC01"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was

Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. .677
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-112N07.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

COMMENT

FEATURES
source

ORIGIN

Query Match 11.8%; Score 146; DB 29; Length 677;
Best Local Similarity 100.0%; Pred. No. 3.3e-61;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 544 GGAGCTCGGTCCTCGTCAGCCATTCAAGCTATTATTGTTCCGGAGCCCGACCT 603
|||||
DB 77 GGAGCTCGGTCCTCGTCAGCCATTCAAGCTATTATTGTTCCGGAGCCCGACCT 136
QY 604 TCAAGTGCAGTGCAGTCTCTGCTGAGCCATTCAAGCTATTATTGTTCCGGAGCCCGACCT 663
|||||
DB 137 TCAAGTGCAGTGCAGTCTCTGCTGAGCCATTCAAGCTATTATTGTTCCGGAGCCCGACCT 196
QY 664 GGCTACCTGGAGAGCTCTGAAGGCC 689
|||||
DB 197 GGCTACCTGGAGAGCTCTGAAGGCC 222

RESULT 11
BX358189/c

LOCUS BX358189 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
DEFINITION clone CSOD1035YC01 3-PRIME, mRNA sequence.
ACCESSION BX358189
VERSION BX358189.1 GI:30370198
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seq@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10299.f
Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CSOD1035AB01NPI.
1. .1201
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1035YC01"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was

FEATURES
source

digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN
Query Match
Best Local Similarity 100.0%; Pred. No. 3.5e-43; Length 1201;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1131 GCCCTCTCCACAGCGGCTGATCAACGTCACCCAGGCGCAGCAAGTCATCCACAGGCTT 1190
Db 811 GCCCTCTCCACAGCGGCTGATCAACGTCACCCAGGCGCAGCAAGTCATCCACAGGCTT 752

QY 1191 CTTTATATCTGCCAGTCTCAGAGGTCAGTAATACAACTGTAGTATAG 1240
Db 751 CTTTATATCTGCCAGTCTCAGAGGTCAGTAATACAACTGTAGTATAG 702

RESULT 12
LOCUS N47851 185 bp mRNA linear EST 14-FEB-1996
DEFINITION YW95h05.r1 Soares_placenta 8to9weeks 2NBHP8to9W Homo sapiens cDNA
ACCESSION N47851
VERSION N47851.1 GI:1189017
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 185)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: T7
High quality sequence stop: 150.
FEATURES
source
1..185
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3889731"
/db_xref="taxon:9606"
/clone="IMAGE:260025"
/dev_stage="two placentae; one from 8 weeks and another from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: placenta; Vector: p77n3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo (dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGCGCAATTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p77n3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN
Query Match
Best Local Similarity 100.0%; Pred. No. 2.2e-38; Length 185;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1141 CCAGCGGCTGATCAACGTCACCCAGGCGCAGCAAGTCATCCACAGGCTTCTTTTAATCT 1200
Db 113 CCAGCGGCTGATCAACGTCACCCAGGCGCAGCAAGTCATCCACAGGCTTCTTTTAATCT 172

QY 1201 GGCCAGTCTCTGAGAAGGTCAGTAATACAACTGTAGTATAG 1240
Db 133 GGCCAGTCTCTGAGAAGGTCAGTAATACAACTGTAGTATAG 172

RESULT 13
LOCUS N93995 234 bp mRNA linear EST 05-APR-1996
DEFINITION za66f09.r1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
ACCESSION N93995
VERSION N93995.1 GI:1266304
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 234)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: mob.REGA+ET
High quality sequence stop: 159.
FEATURES
source
1..234
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1242467"
/db_xref="taxon:9606"
/clone="IMAGE:297545"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: lung; Vector: p77n3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo (dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGCGCAATTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p77n3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal heart library, Soares fetal heart NbHL19W."

ORIGIN
Query Match
Best Local Similarity 100.0%; Pred. No. 2.3e-38; Length 234;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1141 CCAGCGGCTGATCAACGTCACCCAGGCGCAGCAAGTCATCCACAGGCTTCTTTTAATCT 1200
Db 113 CCAGCGGCTGATCAACGTCACCCAGGCGCAGCAAGTCATCCACAGGCTTCTTTTAATCT 172

QY 1201 GGCCAGTCTCTGAGAAGGTCAGTAATACAACTGTAGTATAG 1240
Db 133 GGCCAGTCTCTGAGAAGGTCAGTAATACAACTGTAGTATAG 172

```

Db 173 GCCAGTCTCAGAGAGTCAGTAATACAACTGTAGTATAG 212

RESULT 14
BX396896/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1201 bp mRNA linear EST 13-MAY-2003
BX396896 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI026Y017 3-PRIME, mRNA sequence.
BX396896
BX396896.1 GI:30613250
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10238.f, and
it belongs to a clone representative of this cluster. For more
information about this cluster and the virtual cDNA, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODI026AH09NP1&cluster=10298.f. Contact :
Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CSODI026AH09NP1.
Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI026Y017"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(GT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source
1. .525
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=625 Col=21 Row=E"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites"

ORIGIN
Query Match 6.1%; Score 76; DB 28; Length 525;
Best Local Similarity 100.0%; Pred.No. 2.6e-26;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1165 CAGCCAGCAGTCATCCACAGGCTCTTTTAATCTGGCCAGTCCTCAGAGGTCAGTAA 1224
Db 347 CAGCCAGCAGTCATCCACAGGCTCTTTTAATCTGGCCAGTCCTCAGAGGTCAGTAA 288
Qy 1225 TACAACTGTAGTATAG 1240
Db 287 TACAACTGTAGTATAG 272

Search completed: September 15, 2004, 06:43:45
Job time : 3638.63 secs

Db 173 GCCAGTCTCAGAGAGTCAGTAATACAACTGTAGTATAG 212

RESULT 15
BX396896/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

525 bp DNA linear GSS 13-MAR-1999
HS_5049_A1_C11_T7A RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=625 Col=21 Row=E, genomic survey sequence.
AQ403719
AQ403719
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 525)
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,

Query Match 6.5%; Score 80; DB 13; Length 1201;
Best Local Similarity 100.0%; Pred.No. 3.2e-28;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTGTGAGTCCAGCCACAGTCGTGGATCAGTTTCCTAGGCTGCCATAACAGCACCAT 60
Db 726 AGGTGTGAGTCCAGCCACAGTCGTGGATCAGTTTCCTAGGCTGCCATAACAGCACCAT 667
Qy 61 AACCTGGTGGCTTAGACAA 80
Db 666 AACCTGGTGGCTTAGACAA 647

RESULT 15
AQ403719/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

525 bp DNA linear GSS 13-MAR-1999
HS_5049_A1_C11_T7A RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=625 Col=21 Row=E, genomic survey sequence.
AQ403719
AQ403719
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 525)
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 20:48:25 ; Search time 4828.88 Seconds
(without alignments)
10546.560 Million cell updates/sec

Title: US-09-729-264-1

Perfect score: 1175

Sequence: 1 cgtgtgcccactctgaataa.....gtaatacaactgtagtatag 1175

Scoring table: OLIGO.NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6784142

Minimum DB seq length: 16

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	1175	100.0	1175	6	AX380396	AX380396 Sequence
2	1076	91.6	2051	6	AX747579	AX747579 Sequence
3	1076	91.6	2051	9	AK092516	AK092516 Homo sapi
4	1074	91.4	1240	6	AX380400	AX380400 Sequence
5	1027	87.4	1168	6	AX380398	AX380398 Sequence
6	739	62.9	1139	6	AX380402	AX380402 Sequence
7	302	25.7	142742	9	AF121782	AF121782 Homo sapi
8	302	25.7	340000	9	HS21C080	AL163280 Homo sapi
9	200	17.0	182532	9	CH179X04	AL954228 Pan trogl
10	145	12.3	156288	9	BS000162	BS000162 Pan trogl
11	100	8.5	199665	9	AF064857	AF064857 Homo sapi
12	72	6.1	192219	9	RP43002119	AL954227 Pan trogl
13	53	4.5	40205	9	AF045450	AF045450 Homo sapi
14	53	4.5	170121	9	AF064860	AF064860 Homo sapi
15	25	2.1	6470	3	AB090820	AB090820 Anopheles
16	24	2.0	5178	3	AY119603	AY119603 Drosophila
17	24	2.0	8152	3	DMTOC	Y14157 Drosophila
18	24	2.0	77137	3	AC004422	AC004422 Drosophila
19	24	2.0	83876	2	AC020009	AC020009 Drosophila
20	24	2.0	148102	8	AP003328	AP003328 Oryza sat
21	24	2.0	148762	3	AP002843	AP002843 Oryza sat
22	24	2.0	169931	3	AC008321	AC008321 Drosophila
23	24	2.0	184554	3	AC099022	AC099022 Drosophila
24	24	2.0	268984	3	AE001274	AE001274 Leishmani
25	24	2.0	314957	3	AE003581	AE003581 Drosophila
26	23	2.0	1155	10	AF537215	AF537215 Mus muscu
27	23	2.0	1195	6	AX380404	AX380404 Sequence
28	23	2.0	29865	2	AC020047	AC020047 Drosophila
29	23	2.0	106256	3	AC108135	AC108135 Leishmani
30	23	2.0	124612	2	AC105442	AC105442 Leishmani
31	23	2.0	149964	2	AC120145	AC120145 Mus muscu
32	23	2.0	166000	3	AC104511	AC104511 Drosophila
33	23	2.0	167926	3	AC023685	AC023685 Drosophila
34	23	2.0	181510	2	AC120346	AC120346 Mus muscu
35	23	2.0	186030	8	CNS08C7W	AL731747 Oryza sat
36	23	2.0	196900	2	AC020851	AC020851 Mus muscu
37	23	2.0	332029	3	AE003491	AE003491 Drosophila
38	22	1.9	646	8	BT005060	BT005060 Arabidops
39	22	1.9	757	8	BT004026	BT004026 Arabidops
40	22	1.9	1200	6	A92047	A92047 Sequence 1
41	22	1.9	1200	6	A92049	A92049 Sequence 3
42	22	1.9	1203	6	BD251402	BD251402 Novel neu
43	22	1.9	1203	6	AX658358	AX658358 Sequence
44	22	1.9	1203	9	AF051767	AF051767 Homo sapi
45	22	1.9	1699	6	BD195081	BD195081 Ret ligan

ALIGNMENTS

RESULT 1
AX380396
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AX380396
Sequence 1 from Patent WO0200710.
AX380396.1 GI:19575326
Homo sapiens (human)
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Welcher,A.A., Sarmiento,U.M., Schultz,H.J. and Chute,H.T.
B7-like molecules and uses thereof
Patent: WO 0200710-A 1 03-JAN-2002;

AX380396
Sequence 1 from Patent WO0200710.
AX380396.1 GI:19575326
Homo sapiens (human)
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Welcher,A.A., Sarmiento,U.M., Schultz,H.J. and Chute,H.T.
B7-like molecules and uses thereof
Patent: WO 0200710-A 1 03-JAN-2002;

QY	721	TAGGTTTTCATGTGCTACTCTTCAATACGCTGCTGCTGCTGCGCCCGTTCGTTGTGTGGCT	840
QY	721	TGACGCCGACGTTGACTCTTCAATACGCTGCTGCTGCTGCGCCCGTTCGTTGTGTGGCT	840
QY	781	TGACGCCGACGTTGACTCTTCAATACGCTGCTGCTGCTGCGCCCGTTCGTTGTGTGGCT	840
QY	841	GCAACTGCTGCTGCCGTTTGTGTTTCTGCTGTAGAGAAAAGAGGATTCGTTATTCAT	900
QY	841	GCAACTGCTGCTGCCGTTTGTGTTTCTGCTGTAGAGAAAAGAGGATTCGTTATTCAT	900
QY	901	TTCAAAAGAAATCTGAAAAGAGAGACAAACAAGAACTGACAGAAAGTGGAAATG	960
QY	901	TTCAAAAGAAATCTGAAAAGAGAGACAAACAAGAACTGACAGAAAGTGGAAATG	960
QY	961	AAAACTCCGCTCAATTCAGATGAACAAAGAACCCAGACACCGCTTCTCTCCCTCCCA	1020
QY	961	AAAACTCCGCTCAATTCAGATGAACAAAGAACCCAGACACCGCTTCTCTCCCTCCCA	1020
QY	1021	AATCCTGTGAATCCAGTGAATCCTGAAACAAAGAAACAGTAGCTGTGGCCCTCTCACCAGC	1080
QY	1021	AATCCTGTGAATCCAGTGAATCCTGAAACAAAGAAACAGTAGCTGTGGCCCTCTCACCAGC	1080
QY	1081	GGCTGATCAACTCCACCCAGGCGCAAGTCAATCCACAGGCTTCCTTAAATCTGGCCA	1140
QY	1081	GGCTGATCAACTCCACCCAGGCGCAAGTCAATCCACAGGCTTCCTTAAATCTGGCCA	1140
QY	1141	GTCTGAGAAGGTCAGTAATACAACTGTAGTATAG	1175
QY	1141	GTCTGAGAAGGTCAGTAATACAACTGTAGTATAG	1175
Db	AX747579	Sequence 1104 from Patent EPI308459.	
Db	AX747579	GI:32131967	
REFERENCE			
AUTHORS			
ORGANISM			
SOURCE			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1126; Conservative			
QY	49	ACGGTTCTGGTCTGGTAAATGAAGTCATAGAGGCCCGCCAGNATGCAAGCTCTGAAAGG	108
QY	201	ACGGTTCTGGTCTGGTAAATGAAGTCATAGAGGCCCGCCAGNATGCAAGCTCTGAAAGG	260
QY	109	GCTCCAGGCTCGCTTCACTGCAACCGTCTCCAGGCTTGAAGCTCATCATGTGGGCTC	168
QY	261	GCTCCAGGCTCGCTTCACTGCAACCGTCTCCAGGCTTGAAGCTCATCATGTGGGCTC	320
QY	169	TCAGTGACATGTTGGTGTGAAGCGTTCAGGCCCATGAGGCCCATCATCAAGATGACCGCT	228

321	TCAGTGCATGGTGGTCTAAGCGTCAGGCCCATGGAGCCCATCATCAACCAATGACCCGCT	380
229	TCACCTCTCAGAGGTACGACAGGGCGGGAACCTTCACCTCGAGAGATGATCATCCACAATG	288
381	TCACCTCTCAGAGGTACGACAGGGCGGGAACCTTCACCTCGAGAGATGATCATCCACAATG	440
289	TGGAGCCAGTGATTTCGGGGAAACATCAGATCAGCCTCCAGAACAGTGCCTGCGATGAT	348
441	TGGAGCCAGTGATTTCGGGGAAACATCAGATCAGCCTCCAGAACAGTGCCTGCGATGAT	500
349	CTGTTACTTTACCGTCCAAGTTATGGAGAGCTGTTCAATCCAGTGTAAATCTTTGTAG	408
501	CTGCTTACCTTACCGTCCAAGTTATGGAGAGCTGTTCAATCCAGTGTAAATCTTTGTAG	560
409	TCGCTGAGAATGAACCTTGTGAAGTTACTTGTCTACCTCTACCTGAGACCCGGTCCCGG	468
561	TCGCTGAGAATGAACCTTGTGAAGTTACTTGTCTACCTCTACCTGAGACCCGGTCCCGG	620
469	ATATTTCTCGGAGCTCGGTCTCCTGGTCAGCCATTCAAGCTATTATTTTGTTCGGAGC	528
621	ATATTTCTCGGAGCTCGGTCTCCTGGTCAGCCATTCAAGCTATTATTTTGTTCGGAGC	680
529	CCAGCGACCTTCAAAGTGCATGAGCATCTCTGGTCTTGACCCCAAGCAATGGGACTT	588
681	CCAGCGACCTTCAAAGTGCATGAGCATCTCTGGTCTTGACCCCAAGCAATGGGACTT	740
589	TGACTTCGGTGGCTACCTGGGAAGAGCCTGAAGGCCCGCAAGTCTGCAACTGTAATCTCA	648
741	TGACTTCGGTGGCTACCTGGGAAGAGCCTGAAGGCCCGCAAGTCTGCAACTGTAATCTCA	800
649	CTGTGATTGGTGTCCCCAAGACACTGAGGTGGTATTATTAATTTCCAGGTGATTATCAA	708
801	CTGTGATTGGTGTCCCCAAGACACTGAGGTGGTATTATTAATTTCCAGGTGATTATCAA	860
709	GTTTACCGAGTTTAGTTTTCATTCGCTACTTCGGGCAAGTTGGACTGGACTAGCAG	768
861	GTTTACCGAGTTTAGTTTTCATTCGCTACTTCGGGCAAGTTGGACTGGACTAGCAG	920
769	GCACCATGCTTCTGACGCCGACGTGTACTCTTACAATACGCTGCTGTCGCCGCCGCTC	828
921	GCACCATGCTTCTGACGCCGACGTGTACTCTTACAATACGCTGCTGTCGCCGCCGCTC	980
829	GTTTGTGTGGCTGCAACTGCTGTCGCCGTTGTTGTTTCTGCTGAGAAGAAAAAGAGAT	888
981	GTTTGTGTGGCTGCAACTGCTGTCGCCGTTGTTGTTTCTGCTGAGAAGAAAAAGAGAT	1040
889	TTCGTATTCAATTTCAAAGAAATCTCAAAGAGAGAGACAAACAAGAACTGAGACAG	948
1041	TTCGTATTCAATTTCAAAGAAATCTCAAAGAGAGAGACAAACAAGAACTGAGACAG	1100
949	AAAGTGAAATGAAATCTCCGGCTACAATTCAGATGAACAAAAGACACAGACACCGCTT	1008
1101	AAAGTGAAATGAAATCTCCGGCTACAATTCAGATGAACAAAAGACACAGACACCGCTT	1160
1009	CTCTCCCTCCAAATCTCTGTGAAATCCAGTGATCCTGAAACAAGAAACAGTAGTGGCC	1068
1161	CTCTCCCTCCAAATCTCTGTGAAATCCAGTGATCCTGAAACAAGAAACAGTAGTGGCC	1220
1069	CTCCTCACAGCGGGCTGATCAACGTCACCCAGGCCAGCAAGTCATCCACAGGCTTCTT	1128
1221	CTCCTCACAGCGGGCTGATCAACGTCACCCAGGCCAGCAAGTCATCCACAGGCTTCTT	1280
1129	TTAATCTGGCCAGTCTCTGAGAAGGTCAAGTCAATATCAACTGTAGTATAG	1175
1281	TTAATCTGGCCAGTCTCTGAGAAGGTCAAGTCAATATCAACTGTAGTATAG	1327

RESULT 3					
AK092516					
LOCUS	AK092516	2051 bp	mRNA	linear	PRI 15-JUL-2002
DEFINITION	Homo sapiens	cDNA FLJ35197 fis,	clone PLACE6017788,	highly similar	
				to IGSF5.	

AK092516
 AK092516.1 GI:21751130
 oligo capping; fis (full insert sequence).
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,
 Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
 Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
 Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
 Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
 Oshima, A., Suzuki, Y., Sugano, S., Nagahari, K., Masubo, Y., Nagai, K.
 and Isogai, T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2051)
 Isogai, T. and Yamamoto, J.
 Direct Submission
 Submitted (04-JUL-2002) Takao Isogai, FUJ Project (HRI Team) ; 2-6-7
 Karusa-Kawatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp. Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5' - & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.
 Location/Qualifiers
 1. .2051
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLAC36017788"
 /tissue_type="placenta"
 /clone_lib="PLAC66"
 /note="cloning vector: pME18SFL3"
 FEATURES
 source

Query Match.	91.6%;	Score 1076;	DB 9;	Length 2051;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1126;	Conservative	0;	Mismatches	1; Indels
		0;		Gaps
		0;		
QY	49	ACGGTCTCGGCTCTGGTAATGAAGTCA	TAGAAGCCCCCAGAGTGCACAGTCTCTGAAGG	108
DB	201	ACGGTCTCGGCTCTGGTAATGAAGTCA	TAGAAGCCCCCAGAGTGCACAGTCTCTGAAGG	260
QY	109	GCTCCGAGGCTCGCTTCAAATGCACCGT	TCCGAGGCTCGAAGCTCATATGTGGGCTC	168
DB	261	GCTCCGAGGCTCGCTTCAAATGCACCGT	TCCGAGGCTCGAAGCTCATATGTGGGCTC	320
QY	169	TCAGTGCACATGCTGGTCTAAGCGTTC	GAGGCCCATCAGCCCAATGACCGCT	228
DB	321	TCAGTGCACATGCTGGTCTAAGCGTTC	GAGGCCCATCAGCCCAATGACCGCT	380
QY	229	TCACCTCTCAGAGGTACGACAGCGGGG	GAACCTTCACCTCGAGATGATCATCCACAATG	288
DB	381	TCACCTCTCAGAGGTACGACAGCGGGG	GAACCTTCACCTCGAGATGATCATCCACAATG	440
QY	289	TGAGGCCCATGATTCGGGGACATCAGAT	GCAGCCTCCAGAACAGTCGCTCGATGGAT	348
DB	441	TGAGGCCCATGATTCGGGGACATCAGAT	GCAGCCTCCAGAACAGTCGCTCGATGGAT	500
QY	349	CTGCTTACCTTACCGTCCAACTTATCGG	GAGAGCTGTTCAATCCCACTGTTAACTTTGTAG	408
DB	501	CTGCTTACCTTACCGTCCAACTTATCGG	GAGAGCTGTTCAATCCCACTGTTAACTTTGTAG	560
QY	409	TCGCTGAGATGAACCTTGTGAAGTTACT	TGTCTACCTCAACTGGACCCGGCTCCCGG	468
DB	561	TCGCTGAGATGAACCTTGTGAAGTTACT	TGTCTACCTCAACTGGACCCGGCTCCCGG	620

QY	469	ATATTTCTGGAGCTCGGTCTCTGGTCAGCCATTCAAGCTATTATTTTGTTCGGAGC	528
Db	621	ATATTTCTGGAGCTCGGTCTCTGGTCAGCCATTCAAGCTATTATTTTGTTCGGAGC	680
QY	529	CCAGCGACCTTCAAAAGTGCAGTGAGCATCTCGGTCTGACCCCAACAGCAATGGACTT	588
Db	681	CCAGCGACCTTCAAAAGTGCAGTGAGCATCTCGGTCTGACCCCAACAGCAATGGACTT	740
QY	589	TGACTTGGTGGTCTACCTGGAAGAGCTGAAGCGCCGCAAGTCTGCAACTGTAATCTCA	648
Db	741	TGACTTGGTGGTCTACCTGGAAGAGCTGAAGCGCCGCAAGTCTGCAACTGTAATCTCA	800
QY	649	CTGTGATTCGTCTCCCAAGACACTGAGAGTGGTATTATTAATTCAGAGTATTATCA	708
Db	801	CTGTGATTCGTCTCCCAAGACACTGAGAGTGGTATTATTAATTCAGAGTATTATCA	860
QY	709	GTTTACCGAGTTTAGTGTCTTTCATTCCTACTTGGGGCAAGTTGGACTTGGACTAGCAG	768
Db	861	GTTTACCGAGTTTAGTGTCTTTCATTCCTACTTGGGGCAAGTTGGACTTGGACTAGCAG	920
QY	769	GCACCAATCTTGTGACCGCAGCTGTAATCTTACAATACGCTGCTGCTGCGCGCTC	828
Db	921	GCACCAATCTTGTGACCGCAGCTGTAATCTTACAATACGCTGCTGCTGCGCGCTC	980
QY	829	GTGTGTGGCTGCAACTGCTGCTGCGGTGTGTCTGCTGTAAGAAAGAGGAT	888
Db	981	GTGTGTGGCTGCAACTGCTGCTGCGGTGTGTCTGCTGTAAGAAAGAGGAT	1040
QY	889	TTTCGATTCAATTTCAAAAGAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAG	948
Db	1041	TTTCGATTCAATTTCAAAAGAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAG	1100
QY	949	AAAGTGGAAATGAAATCTCGGCTACAATTCAGATGAACAAAGACACAGACACCGCTT	1008
Db	1101	AAAGTGGAAATGAAATCTCGGCTACAATTCAGATGAACAAAGACACAGACACCGCTT	1160
QY	1009	CTCTCCCTCCCAATCTGTGATTCAGTGCATCTGAAACAAAGAAACAGTAGTGTGGCC	1068
Db	1161	CTCTCCCTCCCAATCTGTGATTCAGTGCATCTGAAACAAAGAAACAGTAGTGTGGCC	1220
QY	1069	CTCTCACACGCGGTGATCAACCTCCACCCAGGCGAGCAAGTCAATCCACAGGCTTCTT	1128
Db	1221	CTCTCACACGCGGTGATCAACCTCCACCCAGGCGAGCAAGTCAATCCACAGGCTTCTT	1280
QY	1129	TTAATCTGGCCAGTCTCTGAGAGGTCAGTAATCAACTGTAGTATAG	1175
Db	1281	TTAATCTGGCCAGTCTCTGAGAGGTCAGTAATCAACTGTAGTATAG	1327
RESULT 4			
AX380400			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1			
Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.			
B7-like molecules and uses thereof			
Patent: WO 0200710-A 5 03-JAN-2002;			
Amgen, Inc. (US)			
FEATURES			
Location/Qualifiers			
1..1240			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
80..1240			
/note="unnamed protein product"			

/codon_start=1		/protein_id="CAD28218.1"	
/db_xref="GI:19575331"		/db_xref="BEMTREMEL:CAD28218"	
/translation="MERHLLTVPERAVGSGSNEVIEGPONATVLKGSQARFNCTVSQG		WKLIMWALSDMVLVSURMEPLIITNDRTSORYDOGNFTSEMIHNYPSPDSGNIRC	
SLQNSRLYFSPDLSQSAVSI1AALTPOQNGTLTCTVATWKSIRKASATVNLTVIRCPQ		VSHGSIYFVPFSPDLSQSAVSI1AALTPOQNGTLTCTVATWKSIRKASATVNLTVIRCPQ	
TGCGGGINPGVLSLSPGLFPGKVLGLAGLMLTPTCLTIRCCCRRCRCCG		NCCRCFCRCRKGFRIFQFKSEKTKETETESNGENSGYNSDEOKTWTASLP	
PKSCSSDPEQRNSSCGPHQADQRP RPASHPOASFNLASPEKVSNTTVV"			
ORIGIN		Query Match	
		Best Local Similarity 91.4%; Score 1074; DB 6; Length 1240;	
		Matches 1124; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	51	GTTCTGGGTCTGGTAATGAAGTCTATAGAAGCCGCCAGAAATGCAACAGTCTCTGAGGCG	110
Db	116	GTTCTGGGTCTGGTAATGAAGTCTATAGAAGCCGCCAGAAATGCAACAGTCTCTGAGGCG	175
QY	111	TCCCAGGCTCGCTCAAACTGCACCGTCTCCAGGCTGGAAGCTCATCATGTGGGCTCTC	170
Db	176	TCCCAGGCTCGCTCAAACTGCACCGTCTCCAGGCTGGAAGCTCATCATGTGGGCTCTC	235
QY	171	AGTGACATCGTGTCTAAAGCGTCAAGCCCATGAGGCCCATCATCAATGACCCGCTTC	230
Db	236	AGTGACATCGTGTCTAAAGCGTCAAGCCCATGAGGCCCATCATCAATGACCCGCTTC	295
QY	231	ACCTCTCAGAGGTACGACAGGCGGGAACCTTCACTCGGAGATGATCATCCAAATGTC	290
Db	296	ACCTCTCAGAGGTACGACAGGCGGGAACCTTCACTCGGAGATGATCATCCAAATGTC	355
QY	291	GAGCCAGATGATTCGGGGAAACATCAGATGAGCGCTCCAGAAAGTCCGCTGATGATCT	350
Db	356	GAGCCAGATGATTCGGGGAAACATCAGATGAGCGCTCCAGAAAGTCCGCTGATGATCT	415
QY	351	GCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTTCAATCCAGTGTATCTTGTAGTC	410
Db	416	GCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTTCAATCCAGTGTATCTTGTAGTC	475
QY	411	GCTGAGATGAACTTGTGAAGTACTTGTCTACCTCTCACCTGACCGGCTCCCGGAT	470
Db	476	GCTGAGATGAACTTGTGAAGTACTTGTCTACCTCTCACCTGACCGGCTCCCGGAT	535
QY	471	ATTTCCTGGAGCTCGGTCTCTGCTGAGCAATCAAGCTATTATTTTGTTCGGAGGCC	530
Db	536	ATTTCCTGGAGCTCGGTCTCTGCTGAGCAATCAAGCTATTATTTTGTTCGGAGGCC	595
QY	531	AGCGACCTTCAAAGTGCAGTGAGCATCTGGGCTCTGACCCCAACAGAGCAATGGACTTTG	590
Db	596	AGCGACCTTCAAAGTGCAGTGAGCATCTGGGCTCTGACCCCAACAGAGCAATGGACTTTG	655
QY	591	ACTTGGTGGCTACCTGGAAGAGCCCTGGAAGCCGCAAGTCTGCACTGTAATCTCACT	650
Db	656	ACTTGGTGGCTACCTGGAAGAGCCCTGGAAGCCGCAAGTCTGCACTGTAATCTCACT	715
QY	651	GTGATTCGGTGTCCCAAGACACTGGAGGTGGTATTATTAATTCAGGTGTATTATCAAGT	710
Db	716	GTGATTCGGTGTCCCAAGACACTGGAGGTGGTATTATTAATTCAGGTGTATTATCAAGT	775
QY	711	TTACCGAGTTTAGGTTTTTTCATTCCTTACCTTACCTGGAAGGCTTGGACTTAGAGGC	770
Db	776	TTACCGAGTTTAGGTTTTTTCATTCCTTACCTTACCTGGAAGGCTTGGACTTAGAGGC	835
QY	771	ACCATGCTTTCAGCGCGAGTGTACTCTTACATACGCTGCTGCTCGCCCGCTCGT	830
Db	836	ACCATGCTTTCAGCGCGAGTGTACTCTTACATACGCTGCTGCTCGCCCGCTCGT	895
QY	831	TGTTGTGGCTGCAACTGCTGCTGCGGTGTGTTTCTGCTGTAGAGAAAAGAGGATTT	890
Db	896	TGTTGTGGCTGCAACTGCTGCTGCGGTGTGTTTCTGCTGTAGAGAAAAGAGGATTT	955

AUTHORS Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
TITLE B7-like molecules and uses thereof
JOURNAL Patent: WO 0200710-A 7 03-JAN-2002;
Amgen, Inc. (US)

FEATURES

Source
1. .1139
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS
1. .1134
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD28219.1"
/db_xref="GI:19575333"
/translation="MVGAMENDDPPGSGSNEVIGPQNAVRLKSGQAFNCTVSQG
WKLIMWALDMVLYSVRPMPEITINDRFTSQRDQGNLTSEMIHNVPSPDSGNIRC
SLONSLHGSAYLTVQVMGELFIPYVVAENEPCEVTLPSHWLRLDISWELGLL
VSHSSYFVPEPDLQSAVSLALTPQSGTLTCVATWKSRLKATKATVNLTVIRCPQ
DTGGGNIYFGLVSLPSLFTMGVGLAGTLMLTPTCTLTTRCCCRRCRCGC
NCCRCGFCRCRRKRNKLRQTKLRQKVEKTPATIQMKRPQTPILLSLPNPNPV
ILNKETIVAVALLTSGLINVHPGQVIRHLLLIWPLVRSVQL"

ORIGIN

Query Match 62.9%; Score 739; DB 6; Length 1139;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 789; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 98 AGTCCTGAAGGGCTCCAGGCTCGCTTCAACTGACCGTCTCCAGGCTGGAAGCTCAT 157
DB 84 AGTCCTGAAGGGCTCCAGGCTCGCTTCAACTGACCGTCTCCAGGCTGGAAGCTCAT 143
QY 158 CATGTGGGCTCTAGTGACATGTTGGTCTTAAGCGTCAGGCGCCATGGAGCCATCATCAC 217
DB 144 CATGTGGGCTCTAGTGACATGTTGGTCTTAAGCGTCAGGCGCCATGGAGCCATCATCAC 203
QY 218 CAATGACCGCTTACCTCTCAGAGTACGACGAGCGGCGGAACTTCACTCGGAGATGAT 277
DB 204 CAATGACCGCTTACCTCTCAGAGTACGACGAGCGGCGGAACTTCACTCGGAGATGAT 263
QY 278 CATCCCAATGTGAGCGCCAGTGTATCGGGGAAACATCAGATCGAGCTCCAGACAGTCG 337
DB 264 CATCCCAATGTGAGCGCCAGTGTATCGGGGAAACATCAGATCGAGCTCCAGACAGTCG 323
QY 338 CCTGATGGATCTGCTTACCTTACCGTCCAGTGTATGGAGAGCTGTTCAATCCAGTGT 397
DB 324 CCTGATGGATCTGCTTACCTTACCGTCCAGTGTATGGAGAGCTGTTCAATCCAGTGT 383
QY 398 TAATCTTGTAGTCTGCTGAGATGAACCTTGTGAAGTTACTTGTCTACCTCAGCTGGAC 457
DB 384 TAATCTTGTAGTCTGCTGAGATGAACCTTGTGAAGTTACTTGTCTACCTCAGCTGGAC 443
QY 458 CCGGCTCCCGGATATTTCTCGGAGCTCGGTCTCTGCTGAGCATTCAAGCTATATTT 517
DB 444 CCGGCTCCCGGATATTTCTCGGAGCTCGGTCTCTGCTGAGCATTCAAGCTATATTT 503
QY 518 TGTTCGGAGCCAGGACCTTCAAGTGCAGTGAAGCTCTGGCTCAGCCCAACAGAG 577
DB 504 TGTTCGGAGCCAGGACCTTCAAGTGCAGTGAAGCTCTGGCTCAGCCCAACAGAG 563
QY 578 CAATGGGACTTTGACTTGGTGGCTTACCTTGGAGAGCTTGAAGCGCCGCAAGCTGCAAC 637
DB 564 CAATGGGACTTTGACTTGGTGGCTTACCTTGGAGAGCTTGAAGCGCCGCAAGCTGCAAC 623
QY 638 TGTAAATCTCATGTGATTCGGTGTCCCAAGACACTGGAGTGGTATTAATATTCAGG 697
DB 624 TGTAAATCTCATGTGATTCGGTGTCCCAAGACACTGGAGTGGTATTAATATTCAGG 683
QY 698 TGTATATCAAGTTTACCGAGTTTAGTGTTCCTTCAATGCTTACTTGGGGCAAGTGGACT 757
DB 684 TGTATATCAAGTTTACCGAGTTTAGTGTTCCTTCAATGCTTACTTGGGGCAAGTGGACT 743
QY 758 TGGACTAGCAGGCACCATCTTCTGACCGCGAGCTGTACTTTTACAATACGCTGTGCTG 817

DB 744 TGGACTAGCAGGCACCATCTTCTGACGCCAGCTGATCTTACAACTAGCTGTGCTG 803
QY 818 CTGCGCCCGCTGTTGTTGGTGGTGAACACTGCTGCTCCCGTGTGTTGTTCTGCTGTAGAAG 877
DB 804 CTGCGCCCGCTGTTGTTGGTGGTGAACACTGCTGCTCCCGTGTGTTGTTCTGCTGTAGAAG 863
QY 878 AAAAAGAGGA 887
DB 864 AAAAAGAGGA 873
RESULT 7
AF121782 142742 bp DNA linear PRI 02-FEB-1999
LOCUS Homo sapiens chromosome 21q22.3 PAC 206A10, complete sequence.
DEFINITION AF121782
ACCESSION AF121782.1 GI:4210991
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 142742)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Taudien, S., Dagand, E., Hildmann, T., Nordsiek, G., Drescher, B.,
Schattevoy, R., Weber, J., Schilling, M., Menzel, U., Yaspo, M.L. and
Rosenthal, A.
TITLE Direct Submission
JOURNAL Submitted (21-JAN-1999) Genome Analysis, Institute for Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES
Location/Qualifiers
1. .142742
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="PAC 206A10"
3. .118
/evidence=not experimental
repeat_region
/rpt_family="MIR"
complement(140..415)
/evidence=not experimental
repeat_region
/rpt_family="MLT2E"
complement(668..881)
/evidence=not experimental
/rpt_family="MER58A"
complement(1078..1162)
/note="GRAIL, score = 86.000%, comment = excellent shadow"
/evidence=not experimental
1114..1403
/evidence=not experimental
/rpt_family="MLT1F"
1481..1603
/evidence=not experimental
/rpt_family="I2"
1576..1710
/note="MZF, score = 66.1%"
/evidence=not experimental
1576..1710
/note="GRAIL, score = 99.000%, comment = excellent"
/evidence=not experimental
complement(1693..1741)
/evidence=not experimental
/rpt_family="MIR"
complement(1780..2326)
/evidence=not experimental
/rpt_family="MLT1F"
complement(1834..1955)
/note="GRAIL, score = 50.000%, comment = good shadow"
/evidence=not experimental
2629..2831
/evidence=not experimental

```

repeat_region /rpt family="AluSx"
3012. .3122 /evidence=not_experimental
repeat_region /rpt family="MLT1J"
3123. .3478 /evidence=not_experimental
repeat_region /rpt family="THE1B"
3479. .3618 /evidence=not_experimental
exon /rpt family="MLT1J"
3502. .3540 /evidence=not_experimental
exon /note="GRAIL, score = 65.000%, comment = good shadow"
4674. .4781 /evidence=not_experimental
exon /note="GRAIL, score = 42.000%, comment = marginal"
4712. .4749 /evidence=not_experimental
repeat_region /note="Xpound exon prediction, score = 70% (0%)"
5385. .6104 /evidence=not_experimental
exon /rpt family="HERV1"
6682. .6751 /evidence=not_experimental
exon /note="MZE, score = 95.1%"
6746. .6751 /evidence=not_experimental
exon /note="Genscan, score = 1.83%, comment = Initial_exon 6 bp
frame: 1 phase: 0"
6901. .7339 /evidence=not_experimental
repeat_region /rpt family="MLT1C"
7347. .7457 /evidence=not_experimental
repeat_region /evidence=not_experimental
rpt family="LRI6C"
7561. .7696 /evidence=not_experimental
repeat_region /rpt family="MIR"
8220. .8220 /evidence=not_experimental
exon /rpt family="MLT1J"
8097. .8119 /evidence=not_experimental
repeat_region /note="GRAIL, score = 45.000%, comment = marginal"
9966. .10041 /evidence=not_experimental
exon /note="homology = 100.00%, score = 38, counts = 2"
repeat_region /rpt_type=tandem
10244. .10664 /rpt unit="ttcttgacacttctgtgtgaggtggtttatcgc"
repeat_region /evidence=not_experimental
rpt family="MLT1G"
10665. .11028 /evidence=not_experimental
repeat_region /rpt family="MLT1F"
11318. .11417 /evidence=not_experimental
repeat_region /rpt family="LIMC2/D"
11503. .11818 /evidence=not_experimental
exon /rpt family="AluSg"
11525. .11777 /evidence=not_experimental
/note="MZE, score = 79.4%"
/evidence=not_experimental
complement (12346. .12827)
/evidence=not_experimental
/rpt family="MLT1C"
12902. .13180 /evidence=not_experimental
/rpt family="AluSg"
complement (13057. .13161)
/note="MZE, score = 76.1%"
/evidence=not_experimental
complement (13226. .13364)
/evidence=not_experimental
/rpt family="LIMC2"
complement (13613. .13908)
/evidence=not_experimental
/rpt family="AluSx"
complement (14317. .14614)
/evidence=not_experimental
/rpt family="LRI37A"
complement (14962. .15036)
/note="MZE, score = 93%"
/evidence=not_experimental
15255. .15299 /evidence=not_experimental
exon /rpt family="LIPa6"
complement (15469. .15497)
/note="Xpound exon prediction, score = 62% (0%)"
16045. .16768 /evidence=not_experimental
/rpt family="MER21A"
complement (16555. .16655)
/note="GRAIL, score = 54.000%, comment = good"
16823. .16850 /evidence=not_experimental
repeat_region /evidence=not_experimental
rpt family="MER34"
complement (16951. .17208)
/evidence=not_experimental
/rpt family="THE1A"
17209. .17632 /evidence=not_experimental
repeat_region /evidence=not_experimental
rpt family="MER34"
complement (17633. .17969)
/evidence=not_experimental
/rpt family="AluSg1"
complement (18004. .18334)
/evidence=not_experimental
/rpt family="LRI17"
complement (18335. .18632)
/evidence=not_experimental
/rpt family="HERV17"
18352. .18420 /evidence=not_experimental
exon /note="GRAIL, score = 51.000%, comment = good shadow"
repeat_region /evidence=not_experimental
complement (18397. .18594)
/note="GRAIL, score = 66.000%, comment = good"
repeat_region /evidence=not_experimental
complement (18643. .24647)
/evidence=not_experimental
/rpt family="HERV17"
complement (19228. .19361)
exon /note="MZE, score = 92.8%"
/evidence=not_experimental
20900. .21135 /evidence=not_experimental
/note="MZE, score = 51.4%"
/evidence=not_experimental
21150. .21275 /evidence=not_experimental
/note="Xpound exon prediction, score = 74% (0%)"

```

Query Match 25.7%; Score 302; DB 9; Length 142742;
 Best Local Similarity 100.0%; Pred. No. 1.6e-166;
 Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 368 AGTTATGGAGAGCTGTTTCATCCAGTGTAAATCTTGTAGTCGCTGAGAAATGAACCTTG 427
Db 49654 AGTTATGGAGAGCTGTTTCATCCAGTGTAAATCTTGTAGTCGCTGAGAAATGAACCTTG 49713
QY 428 TGAAGTTACTTGTCTACCTTCACACTGGACCGGCTCCCGGATATTTCTCGGAGCTCGG 487
Db 49714 TGAAGTTACTTGTCTACCTTCACACTGGACCGGCTCCCGGATATTTCTCGGAGCTCGG 49773
QY 488 TCTCTGGTCAGCAATCAAGCTATTATTTGTTCCGGAGCCAGGACCTTCAAAGTGC 547
Db 49774 TCTCTGGTCAGCAATCAAGCTATTATTTGTTCCGGAGCCAGGACCTTCAAAGTGC 49833
QY 548 AGTGAGCATCTGGCTCTGACCCACAGACCAATGGACTTTGACTTGGTGGCTACCTTG 607
Db 49834 AGTGAGCATCTGGCTCTGACCCACAGACCAATGGACTTTGACTTGGTGGCTACCTTG 49893
QY 608 GAAGAGCCTGAAGCGCCGCAAGTCTGCAACTGTAATCTCACTGTGATTCGGTGTCCCA 667
Db 49894 GAAGAGCCTGAAGCGCCGCAAGTCTGCAACTGTAATCTCACTGTGATTCGGTGTCCCA 49953
QY 668 AG 669
Db 49954 AG 49955
RESULT 8
HS21C080 340000 bp DNA linear PRI 24-MAY-2000
LOCUS Homo sapiens chromosome 21 segment HS21C080.
DEFINITION AL163280 AP001735 BA000005
ACCESSION AL163280.2 GI:7717369
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 340000)
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
Okai,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
Aakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuayama,S.,
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,
Hornischer,K., Brandt,P., Scharfe,M., Schoen,O., Desario,A.,
Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S.,
Hennig,S., Rieselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K.,
Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
Yaspo,M.L.
Direct Submission
Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing
Consortium: * RIKEN Genomic Sciences Center, Human Genome Research
Group * Institute of Molecular Biotechnology, Genome Analysis *
Keio University School of Medicine, Dept. of Molecular Biology *
GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular
Genetics (addresses see below)
The Chromosome 21 Mapping and Sequencing Consortium consists of
* RIKEN Genomic Sciences Center, Human Genome Research Group, *
Sagamihara 228-8555, Japan,
* e.mail: sakaki@gs.c.riken.go.jp
* URL: http://hgp.gsc.riken.go.jp/
and
* Institute of Molecular Biotechnology, Genome Analysis, *
Beutenbergstrasse 11, D-07745 Jena, Germany,
* e.mail: gscj-submit@genome.imb-jena.de
* URL: http://genome.imb-jena.de/
and
* Keio University School of Medicine, Dept. of Molecular Biology, *
Tokyo 160-8582, Japan,
* e.mail: shimizudmb-med.keio.ac.jp
* URL: http://adenine.dmb.med.keio.ac.jp/
and

* GBF, Dept. of Genome Analysis,
* Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail:
info.genome@gbf.de
* URL: http://genome.gbf.de/
and
* Max-planck Institute for Molecular Genetics,
* Inneustraße 73, D-14195 Berlin, Germany,
* e.mail: info-chr21@molgen.mpg.de
* URL: http://chr21.rz-berlin.mpg.de/.
Location/Qualifiers
1. 340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
<1. 125946
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P160L9, 5' partial"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AF121897"
73410..243533
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P70124"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AF064860"
224137..>340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P206A10, 3' partial"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AF121782"
286628..>340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="BAC-291B3, 3' partial"
/clone_lib="BAC library"
/note="Accession No. AF064857"
1088..1566
/note="L1M4"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
2398..2495
/note="(TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(2496..2878)
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
2879..2931
/note="(TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
2932..3004
/note="(TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3005..3074
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region

```

/note="(TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3473..19157
/gene="SH3BGR"
/mRNA
join<3473..3565,12649..12678,15331..15460,18739..19157)
/gene="SH3BGR"
CDS
join(3473..3565,12649..12678,15331..15426)
/gene="SH3BGR"
/note="Accession No. X93498"
/codon_start=1
/product="21-Glutamic Acid Rich protein 21-GARP"
/protein_id="CAB90445.1"
/db_xref="GI:7717370"
/db_xref="GOA:P55822"
/db_xref="SWISS-PROT:P55822"
/translation="GSEKAEGGTEAQKGESEDVGNLPFAQKNEBERGETATETETE
IAMSGAGAEAEETARGETEPGEDEDS"
3473..3565
/exon
/gene="SH3BGR"
/number=4
repeat_region
5542..5829
/note="AluX"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(6887..7067)
/note="L1MD3"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
complement(8482..8762)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(8865..8964)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
9131..9424
/note="AluSq"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(9723..9882)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(9989..10077)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(10078..10429)
/note="THE1B"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
complement(10430..10651)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
10980..11000
/note="(TTTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(11002..11282)
/note="AluSp"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(11315..11556)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
11702..11753
/note="MIR"
/rpt_family="SINE/MIR"
/rpt_type=DISPERSED

```

```

exon      12649..12678
          /gene="SH3BGR"
          /number=5
repeat_region
          complement(13115..13425)
          /note="AluX"
          /rpt_family="SINE/Alu"
          /rpt_type=DISPERSED
repeat_region
          complement(13598..13892)

Query Match      25.7%   Score 302;   DB 9;   Length 340000;
Best Local Similarity 100.0%;   Pred. No. 1.7e-166;
Matches 302;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY 368 AGTTATGGAGAGCTGTTTCATCCAGTGTAACTTTAGTCGCTGAGATGAACCTTG 427
Db 273834 AGTTATGGAGAGCTGTTTCATCCAGTGTAACTTTAGTCGCTGAGATGAACCTTG 273893

QY 428 TGAAGTTACTTGTCTACCTCCACACTGGACCCGGCTCCCGGATATTTCTCGGAGCTCGG 487
Db 273894 TGAAGTTACTTGTCTACCTCCACACTGGACCCGGCTCCCGGATATTTCTCGGAGCTCGG 273953

QY 488 TCTCTGTGTAGCCATTCAGCTATTATTTTGTTCGGAGGCCAGCGACCTTCAAAGTGC 547
Db 273954 TCTCTGTGTAGCCATTCAGCTATTATTTTGTTCGGAGGCCAGCGACCTTCAAAGTGC 274013

QY 548 AGTGAGCATCTGGCTCTGACCCACAGAGCAATGGGACTTTTGACTTGCCTGGCTACCTG 607
Db 274014 AGTGAGCATCTGGCTCTGACCCACAGAGCAATGGGACTTTTGACTTGCCTGGCTACCTG 274073

QY 608 GAAGAGCTTGAAGGCCCGCAAGTCTGCAACTGTAAATCTCACTGTGATTGGTGTCCCCA 667
Db 274074 GAAGAGCTTGAAGGCCCGCAAGTCTGCAACTGTAAATCTCACTGTGATTGGTGTCCCCA 274133

QY 668 AG 669
Db 274134 AG 274135

RESULT 9
CH179K04      182532 bp   DNA   linear   PRI 16-OCT-2003
LOCUS      Pan troglodytes chromosome 22 BAC CH251-179K04, complete sequence.
DEFINITION
ACCESSION   AL954228
VERSION     AL954228.1 GI:37619870
KEYWORDS    HTG.
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM    Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1 (bases 1 to 182532)
            The Chimpanzee Chromosome 22 Sequencing Consortium
            Chimpanzee chromosome 22 genomic sequence
            Unpublished
REFERENCE   2 (bases 1 to 182532)
            Scharfe M., Berg C., Conrad A., Hornischer K., Loehnert T.H.,
            Ludewig M., Thies S., Weber K. and Bloeker H.
            Direct Submission
            Submitted (03-JUN-2003) GBF, Dept. of Genome Analysis, Mascheroder
            Weg 1, D-38124 Braunschweig, Germany, E-mail: info.genome@gbf.de
            The Chimpanzee Chromosome 22 Sequencing Consortium consists of :
            Shanghai, China
            Taiwan;
            ----- Genome Center
            Center: GBF, Braunschweig
            Center code: GBF
            Web site: http://genome.gbf.de/
            Contact: info.genome@gbf.de
            ----- Project information
            Center project name:
            Center clone name: CH251-179K04
            ----- Summary Statistics
            Sequencing vector: ###;

```

```

Chemistry: Dye-terminator-amersham: ## of reads
Chemistry: Dye-primer-amersham: ## of reads
Assembly program: ##
Consensus quality: 0 bases at least Q40
Consensus quality: 0 bases at least Q30
Consensus quality: 0 bases at least Q20
Estimated insert size: #: agarose-fp estimation
Estimated insert size: 182532; sum-of-contigs estimation
-----
PROGRAMS AND PARAMETERS USED FOR ANNOTATION:
+*****
+ Analysis and annotation were performed with the automatic
+ 'first-pass' annotation and submission tool
+ 'AnnoMitter' (Hornischer & Bloeker).
+ Programs used by 'AnnoMitter':
+*****
Location/Qualifiers
1. 182532
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="22"
/clone="CH251-179K04"
1. 182532
/note="assembly fragment-clone end:T7-vector side:left
assembly_fragment-clone end:SP6-vector side:right"

Query Match 17.0%; Score 200; DB 9; Length 182532;
Best Local Similarity 100.0%; Pred. No. 5.4e-106;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 460 GGCTCCGGATATTTCTGGAGTCGGTCTCTGTCAGCCATTCAGCTATTATTG 519
Db 17884 GGCTCCGGATATTTCTGGAGTCGGTCTCTGTCAGCCATTCAGCTATTATTG 17943

QY 520 TTCGGAGCCAGGACCTTCAAGTCAGTGACATCTGGCTGACCCACAGAGCA 579
Db 17944 TTCGGAGCCAGGACCTTCAAGTCAGTGACATCTGGCTGACCCACAGAGCA 18003

QY 580 ATGGGACTTTGACTTGGTGCTACCTGGAAGACCTGGAAGCCCGGAGTCTCAACTG 639
Db 18004 ATGGGACTTTGACTTGGTGCTACCTGGAAGACCTGGAAGCCCGGAGTCTCAACTG 18063

QY 640 TAAATCTCACTGTGATTCGG 659
Db 18064 TAAATCTCACTGTGATTCGG 18083

RESULT 10
BS000162 156288 bp DNA linear PRI 07-OCT-2003
LOCUS Pan troglodytes chromosome 22 clone:PTB-060F12, map 22, complete
DEFINITION sequences.
ACCESSION BS000162
VERSION BS000162.1 GI:37537429
KEYWORDS HTG.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE
1 The Chimpanzee Chromosome 22 Sequencing Consortium.
DNA sequence of chimpanzee chromosome 22 and its evolutionary
implications
Unpublished
JOURNAL
2 (bases 1 to 156288)
AUTHORS Saitou,N., Kim,C., Kitano,T., Oota,S., Shimada,M., Kryukov,K.,
Tomiki,T. and Kohara,Y.
REFERENCE
Submitted (15-MAY-2003) Naruya Saitou, National Institute of
Genetics (NIG), Division of Population Genetics, Illi Yata,
Mishima, Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,

COMMENT
URL:http://sayer.lab.nig.ac.jp/, Tel:81-55-981-6790,
Fax:81-55-981-6789)
The Chimpanzee Chromosome 22 Sequencing Consortium consists of:
*Chinese National Human Genome Center at Shanghai, Shanghai, China;
*GBF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
of Molecular Biotechnology, Jena, Germany; *KIBB Genome Research
Center, Daejeon, Korea;
*Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
*National Institute of Genetics, Mishima, Japan;
*National Yang Ming University Genome Research Center, Taipei,
Taiwan;
*RIKEN Genomic Sciences Center, Yokohama, Japan.
----- Genome Center
Center: National Institute of Genetics
Center code: NIG
Web site: http://sayer.lab.nig.ac.jp/
Contact: nsaitou@genes.nig.ac.jp
----- Project Information
Center project name: The Chimpanzee Chromosome 22 Sequencing Project
Center clone name: PTB-060F12
----- Summary Statistics
Sequencing vector: pUC118; 100% of reads
Chemistry: Dye-terminator Big Dye and dGTP; 100% of reads Assembly
program: Phrap; version 0.990329
Consensus quality: 156,231 bases at least Q40
Consensus quality: 156,288 bases at least Q30
Consensus quality: 156,288 bases at least Q20
-----
This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30);
an attempt was made to resolve all sequencing problems, such as
compressions and repeats; all regions were covered by at one
plasmid
subclone or more than one M13 subclone;
-----
Source information:
The PTB1 chimpanzee BAC library was prepared from DNA isolated from
cultured cells established from the blood of a single male
chimpanzee.
Clones may be obtained from Asao Fujiyama and co-workers
(http://www.gsc.riken.go.jp).
VECTOR: pKS145
-----
Sequence Quality Assessment:
This entry has been annotated with sequence
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in
10,000 bp.
-----
Neighboring clones: CH251-179K04 (left) and PTB-103H04 (right).
-----
Location/Qualifiers
1. 156288
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="22"
/clone="PTB-060F12"
/clone_lib="PTB1 chimpanzee BAC"

FEATURES
source
Query Match 12.3%; Score 145; DB 9; Length 156288;
Best Local Similarity 100.0%; Pred. No. 2.3e-73;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
QY 480 GAGTCGGTCTCTGTCAGCCATTCAGCTATTATTGTCGCGCCAGCCGACCTT 539
Db 1 GAGTCGGTCTCTGTCAGCCATTCAGCTATTATTGTCGCGCCAGCCGACCTT 60
QY 540 CAAGTCGAGTCAGCATCTGCTGACCCACAGAGCAATGGGACTTTGACTTCGCTG 599

```

```

Db      61 CAAAGTCAGTCAGCATCTGCTCTGACCCACAGAGCAATGGGACTTTTGACTTGCCTG 120
QY      600 GCTACTGGAAGAGCTGAGGCC 624
Db      121 GCTACTGGAAGAGCTGAGGCC 145

RESULT 11
AF064857
LOCUS   AF064857                199665 bp    DNA    linear    PRI 23-JUN-1998
DEFINITION Homo sapiens BAC derived from chromosome 21q22.3, complete
sequence, containing PEP19 (PCP4) gene.
ACCESSION AF064857
VERSION   AF064857.1   GI:3171149
KEYWORDS HTG.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 199665)
Taudien,S., Nordliek,G., Korenberg,J., Drescher,B., Weber,J.,
Schatnevoy,R. and Rosenthal,A.
Submitted (11-MAY-1998) Genome Analysis, Institute for Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES             Location/Qualifiers
     source            1..199665
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.3"
     complement(1..383)
     rpt_family="L1"
     repeat_region     747..814
                     /evidence=not_experimental
     repeat_region     747..814
                     /evidence=not_experimental
     rpt_family="MIR"
     repeat_region     complement(1345..1500)
                     /evidence=not_experimental
     rpt_family="L1PA4"
     repeat_region     1499..1588
                     /evidence=not_experimental
     rpt_family="L1PA2"
     repeat_region     complement(1584..2179)
                     /evidence=not_experimental
     rpt_family="L1PA2"
     repeat_region     2174..3405
                     /evidence=not_experimental
     rpt_family="L1PA2"
     repeat_region     3388..3454
                     /evidence=not_experimental
     rpt_family="L1P"
     repeat_region     3968..4086
                     /evidence=not_experimental
     rpt_family="MIR"
     repeat_region     complement(6245..6668)
                     /evidence=not_experimental
     rpt_family="L1ME3A"
     repeat_region     6671..6969
                     /evidence=not_experimental
     rpt_family="AluJb"
     repeat_region     complement(7057..7376)
                     /evidence=not_experimental
     rpt_family="MER2"
     repeat_region     complement(7387..7491)
                     /evidence=not_experimental
     rpt_family="L1ME3A"
     repeat_region     7781..8020
                     /evidence=not_experimental
     rpt_family="MER58A"
     repeat_region     8650..8823
                     /evidence=not_experimental

/rpt_family="L1PA5"
complement(10456..10720)
/evidence=not_experimental
/rpt_family="L2"
complement(13514..13814)
/evidence=not_experimental
/rpt_family="AluSx"
complement(13823..14116)
/evidence=not_experimental
/rpt_family="AluJo"
complement(15662..15785)
/evidence=not_experimental
/rpt_family="MER5A"
complement(18994..19347)
/evidence=not_experimental
/rpt_family="THE1B"
19378..19470
/note="homology = 81.70%, score = 28, counts = 3"
/evidence=not_experimental
/rpt_type=tandem
/rpt_unit="agtatctattctgattggcgagtcgtcatcac"
complement(20371..20556)
/evidence=not_experimental
/rpt_family="MER3"
21213..21500
/evidence=not_experimental
/rpt_family="AluSx"
22291..22567
/evidence=not_experimental
/rpt_family="MER39b"
complement(22583..23201)
/evidence=not_experimental
/rpt_family="L1PA2"
23202..23471
/evidence=not_experimental
/rpt_family="MER39b"
24223..24454
/evidence=not_experimental
/rpt_family="MSTB-internal"
24455..24889
/evidence=not_experimental
/rpt_family="MSTB"
complement(25434..25767)
/evidence=not_experimental
/rpt_family="MER58B"
25993..26294
/evidence=not_experimental
/rpt_family="AluSq"
27222..27283
/evidence=not_experimental
/rpt_family="MIR"
complement(28932..29335)
/evidence=not_experimental
/rpt_family="MSTA"
complement(29462..30001)
/evidence=not_experimental
/rpt_family="MER34"
complement(30308..30676)
/evidence=not_experimental
/rpt_family="MLT1B"
31012..31364
/evidence=not_experimental
/rpt_family="MLT1A2"
complement(32507..32738)
/evidence=not_experimental
/rpt_family="MSTD"
complement(32766..32980)
/evidence=not_experimental
/rpt_family="MSTD"
33468..33678
/evidence=not_experimental
/rpt_family="THE1B"
34563..34620

```

/note="homology = 93.10%, score = 21, counts = 2"
/evidence=not experimental
/rpt type=tandem
/rpt unit="tttccagaagaatacgtacttaggca"
complement(35785. .35824)
/evidence=not experimental
/rpt family="MLT2CB"
complement(35924. .36345)
/evidence=not experimental
/rpt family="MLT2CB"
complement(36403. .36475)
/evidence=not experimental
/rpt family="MLT2CB"
complement(36717. .36828)
/evidence=not experimental
/rpt family="MER67C"
complement(37064. .37364)
/evidence=not experimental
/rpt family="AluY"
37440. .37501
/note="homology = 98.40%, score = 29, counts = 2"
/evidence=not experimental
/rpt type=tandem
/rpt unit="aaaacactcacttcttcaaaactgctc"
37518. .38158
/evidence=not experimental
/rpt family="L2"
38477. .39376
/evidence=not experimental
/rpt family="L2"
39552. .39861
/evidence=not experimental
/rpt family="MLTID"
39884. .40122
/evidence=not experimental
/rpt family="MLTID"
complement(40243. .40542)
/evidence=not experimental
/rpt family="AluSx"
40543. .41180
/evidence=not experimental
/rpt family="L2"
complement(42066. .42114)
/evidence=not experimental
/rpt family="MIR"
42138. .42317
/evidence=not experimental
/rpt family="L2"
complement(42332. .42710)
/evidence=not experimental
/rpt family="L1P81"
complement(42758. .46751)
/evidence=not experimental
/rpt family="L1P81"
complement(46792. .47084)
/evidence=not experimental
/rpt family="L1MA"
47189. .47992
/evidence=not experimental
/rpt family="L2"
complement(48002. .48132)
/evidence=not experimental
/rpt family="MIR"

Query Match 8.5%; Score 100; DB 9; Length 199665;
Best Local Similarity 100.0%; Pred. No. 1.1e-46;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1076 CCAGCGGCTGTACAGTCCACCCAGCAGTATCCACAGGCTTCTTTAACTC 1135
|||||
DB 17555 CCAGCGGCTGTACAGTCCACCCAGCAGTATCCACAGGCTTCTTTAACTC 17614
QY 1136 GCCCAGTCTGAGAGGTCAGTAATACAACTGTAGTATAG 1175

Db 17615 GCCCAGTCTGAGAGGTCAGTAATACAACTGTAGTATAG 17654
|||||
RESULT 12
RP43002119 192219 bp DNA linear PRI 19-NOV-2003
LOCUS Pan troglodytes chromosome 22 BAC RP43-002119, complete sequence.
DEFINITION AL954227
VERSION AL954227.3 GI:38453654
KEYWORDS HTG.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM
REFERENCE
AUTHORS Pan troglodytes
TITLE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL Eukaryota; Metazoa; Primates; Catarrhini; Hominidae; Pan.
AUTHORS
REFERENCE 1 (bases 1 to 192219)
TITLE The Chimpanzee Chromosome 22 Sequencing Consortium.
JOURNAL Chimpanzee chromosome 22 genomic sequence
AUTHORS Unpublished
REFERENCE 2 (bases 1 to 192219)
TITLE Schafie, M., Berg, C., Conrad, A., Hornischer, K., Loehnert, T.H.,
JOURNAL Ludwig, M., Thies, S., Weber, K. and Bloecker, H.
AUTHORS Direct Submission
TITLE Submitted (25-JUN-2003) GBF, Dept. of Genome Analysis, Mascheroder
JOURNAL Weg 1, D-38124 Braunschweig, Germany, E-mail: info.genome@gbf.de
COMMENT On Nov 19, 2003 this sequence version replaced gi:38228900.
The Chimpanzee Chromosome 22 Sequencing Consortium consists of :
*Chinese National Human Genome Center at Shanghai,
Shanghai, China
*GBF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
of Molecular Biotechnology, Jena, Germany; *KRIBB Genome Research
Center, Daejeon, Korea;
*Max-Planck Institute for Molecular Genetics, Berlin, Germany;
*National Institute of Genetics, Mishima, Japan;
*National Yang Ming University Genome Research Center, Taipei,
Taiwan;
*RIKEN Genomic Sciences Center, Yokohama, Japan.
PROGRAMS AND PARAMETERS USED FOR ANNOTATION:
+ Analysis and annotation were performed with the automatic
+ 'first-pass' annotation and submission tool
+ 'AnnoMitter' (Hornischer & Bloecker).
+ Programs used by 'AnnoMitter':
+----- Genome Center
Center: GBF, Braunschweig
Center code: GBF
Web site: http://genome.gbf.de/
Contact: info.genome@gbf.de
----- Project Information
Center project name:
Center clone name: RP43-002119
----- Summary Statistics
Sequencing vector: ##;
Chemistry: Dye-terminator-amersham: ## of reads
Chemistry: Dye-primer-amersham: ## of reads
Assembly program: #
Consensus quality: 0 bases at least Q40
Consensus quality: 0 bases at least Q30
Consensus quality: 0 bases at least Q20
Estimated insert size: ##; agarose-fp estimation
Estimated insert size: 192219; sum-of-contigs estimation

* Location/Qualifiers
1. 192219
/organism="pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="22"
/clone="RP43-002119"
1. 192219
/note="assembly_fragment-clone_end:T7-vector_side:left
misc_feature
FEATURES
Source

```

ORIGIN      assembly_fragment-clone_end.SP6-vector_side::right"

Query Match      6.1%; Score 72; DB 9; Length 192219;
Best Local Similarity 100.0%; Pred. No. 4.6e-30;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 GCGAAGTTCACCTCGAGATGATCATCCCAATGTGGAGCCGAGTCGATTCGGGGAAACATC 314
      |||||
Db 191622 GCGAAGTTCACCTCGAGATGATCATCCCAATGTGGAGCCGAGTCGATTCGGGGAAACATC 191681

QY 315 AGATGACGAGCTC 326
      |||||
Db 191682 AGATGACGAGCTC 191693

RESULT 13
LOCUS      AF045450          40205 bp      DNA      linear      PRI 20-MAR-1998
DEFINITION Homo sapiens chromosome 21q22.3 cosmid Q11M15, complete sequence.
ACCESSION  AF045450
VERSION    AF045450.1 GI:2895783
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 40205)
AUTHORS   Taudien,S. and Rosenthal,A.
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 40205)
AUTHORS   Taudien,S., Nordieck,G., Dagand,E., Hildmann,T., Drescher,B.,
            Weber,J., Rosenthal,A. and Yaspo,M.L.
TITLE     Direct Submission
JOURNAL   Submitted (29-JAN-1998) Genome Analysis, Institut for Molecular
            Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES   Location/Qualifiers
            1..40205
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="21"
                /map="21q22.3"
                /clone="cosmid Q11M15"
                /complement(293..586)
                /evidence=not_experimental
                /rpt_family="AluSc"
                /complement(643..964)
                /evidence=not_experimental
                /rpt_family="MERVL"
                /complement(1053..1176)
                /notes="GenScan, score = 5.19%, comment = Internal_exon 124
                bp frame: 2 phase: 1"
                /evidence=not_experimental
                /complement(1251..1859)
                /evidence=not_experimental
                /rpt_family="HERVL"
                /complement(2296..3220)
                /evidence=not_experimental
                /rpt_family="HERV16"
                /complement(2818..2889)
                /note="Xpound exon prediction, score = 82% (0%)"
                /evidence=not_experimental
                /complement(3221..3658)
                /evidence=not_experimental
                /rpt_family="MLT1C"
                /complement(3359..3395)
                /note="Xpound exon prediction, score = 84% (0%)"
                /evidence=not_experimental
                /complement(3874..4176)
                /evidence=not_experimental
                /rpt_family="AluJb"
                /complement(4330..4433)

```

```

/evidence=not_experimental
/rpt_family="L2"
4585..4638
/note="GRAIL, score = 95.000%, comment = excellent shadow"
/evidence=not_experimental
complement(5113..7899)
/evidence=not_experimental
/rpt_family="L1PA14"
complement(6119..6598)
/note="GRAIL, score = 63.000%, comment = good"
/evidence=not_experimental
complement(7927..8045)
/evidence=not_experimental
/rpt_family="FLAM_A"
8048..8178
/evidence=not_experimental
/rpt_family="L2"
8771..8876
/note="MZEF, score = 67.4%"
/evidence=not_experimental
complement(8809..9187)
/evidence=not_experimental
/rpt_family="MER1A"
8837..8876
/note="GRAIL, score = 55.000%, comment = good shadow"
/evidence=not_experimental
complement(9113..9213)
/note="GRAIL, score = 55.000%, comment = good"
/evidence=not_experimental
complement(9190..9258)
/evidence=not_experimental
/rpt_family="L1MC3"
9259..9647
/evidence=not_experimental
/rpt_family="L1MC3"
complement(10138..10427)
/evidence=not_experimental
/rpt_family="MLT2E"
10662..10875
/evidence=not_experimental
/rpt_family="MER58A"
complement(11072..11156)
/note="GRAIL, score = 86.000%, comment = excellent"
/evidence=not_experimental
complement(11072..11129)
/note="Xpound exon prediction, score = 89% (0%)"
/evidence=not_experimental
11478..11597
/evidence=not_experimental
/rpt_family="L2"
11570..11704
/note="MZEF, score = 66.1%"
/evidence=not_experimental
11570..11704
/note="GRAIL, score = 99.000%, comment = excellent shadow"
/evidence=not_experimental
complement(11687..11735)
/evidence=not_experimental
/rpt_family="Mir"
complement(11774..12320)
/evidence=not_experimental
/rpt_family="MLT1E"
complement(11828..11949)
/note="GRAIL, score = 50.000%, comment = good"
/evidence=not_experimental
12623..12843
/evidence=not_experimental
/rpt_family="AluSx"
13102..13457
/evidence=not_experimental
/rpt_family="THE1B"
complement(13481..13519)
/note="GRAIL, score = 65.000%, comment = good"

```

/evidence=not_experimental
15065. .15188
/evidence=not_experimental
/rpt family="HERVL"
15130. .15378
/note="GC score = 10.60 (249bp); Region: GC content"
/evidence=not_experimental
15369. .16268
/evidence=not_experimental
/rpt family="MERVL"
complement(16263. .16714)
/note="GRAIL, score = 72.000%, comment = good"
/evidence=not_experimental
16661. .16730
/note="WZEF, score = 93.5%"
/evidence=not_experimental
16880. .17318
/evidence=not_experimental
/rpt family="MLT1C"
17326. .17436
/evidence=not_experimental
/rpt family="LTRL1C"
complement(17540. .17675)
/evidence=not_experimental
/rpt family="MIR"
complement(17805. .17864)
/evidence=not_experimental
/rpt family="MLT1C"
18076. .18098
/note="GRAIL, score = 45.000%, comment = marginal shadow"
/evidence=not_experimental
complement(18219. .18349)
/note="Genscan, score = 2.99%, comment = Internal_exon 131
bp frame: 1 phase: 2"
/evidence=not_experimental
complement(18717. .19211)
/evidence=not_experimental
/rpt family="MLT1C"
complement(19252. .19399)
/note="GRAIL, score = 71.000%, comment = good"
/evidence=not_experimental
19269. .19350
/note="GRAIL, score = 81.000%, comment = excellent shadow"
/evidence=not_experimental
complement(19379. .19399)
/note="Xpound exon prediction, score = 60% (0%)
/evidence=not_experimental
complement(19501. .19881)
/note="WZEF, score = 50.6%"
/evidence=not_experimental
19945. .20020
/note="homology = 100.00%, score = 38, counts = 2"
/evidence=not_experimental
/rpt type=tandem
/rpt unit="ttcttcacactctctgtgagaggtcggtttatgcg"
20223. .20643
/evidence=not_experimental
/rpt family="MLT1B"
20645. .20910
/evidence=not_experimental
/rpt family="MLT1F"
20961. .21007
/evidence=not_experimental
/rpt family="MLT1F"
21483. .21781
/evidence=not_experimental
/rpt family="AluJb"
complement(21504. .21756)
/note="WZEF, score = 79.4%"
/evidence=not_experimental
complement(22325. .22806)
/evidence=not_experimental

Query Match 4.5%; Score 53; DB 9; Length 40205;
Best Local Similarity 100.0%; Pred. No. 8e-19; Indels 0; Gaps 0;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGTCTGCCATCTCAATAACAAGAGATGGGCTTGTGATTTTCTCCACGGT 53
DB 14773 CTGTCTGCCATCTCAATAACAAGAGATGGGCTTGTGATTTTCTCCACGGT 14825
RESULT 14
AF064860 170121 bp DNA linear PRI 05-MAR-2002
LOCUS Homo sapiens chromosome 21 clone PAC 70124 map 21q22.3, complete
DEFINITION sequence.
ACCESSION AF064860
VERSION AF064860.2 GI:18958624
KEYWORDS HTG; HTGS DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 170121)
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
Park,H.-S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.-K., Soeda,E.,
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
Patterson,D., Reichwald,K., Rump,A., Schillhabel,M.B., Schudy,A.,
Zimmermann,W., Rosenthal,A., Kudoh,J., Kawasaka,S.,
Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S.,
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,
Hornischer,K., Brandt,P., Scharfe,M., Schoen,O., Desario,A.,
Reichelt,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S.,
Hennig,S., Rieselmann,L., Dagand,E., Haaf,T., Wehrmeyer,S.,
Borzym,K., Gardiner,K., Nizetic,D., Francis,F., Lehrach,H.,
Reinhardt,R. and Yaspo,M.Laure.
The DNA sequence of human chromosome 21
Nature 405 (6784), 311-319 (2000)
20289799
10830953
REFERENCE 2 (bases 1 to 170121)
Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.
Direct Submission
Submitted (12-MAY-1998) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
REFERENCE 3 (bases 1 to 170121)
Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.
Rosenthal,A.
Direct Submission
Submitted (27-FEB-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
REFERENCE 4 (bases 1 to 170121)
Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.
Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.
Direct Submission
Submitted (05-MAR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
On Feb 27, 2002 this sequence version replaced gi:171153.
FEATURES
Location/Qualifiers
1..170121
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="PAC 70124"

ORIGIN
Query Match 4.5%; Score 53; DB 9; Length 170121;
Best Local Similarity 100.0%; Pred. No. 8.6e-19; Indels 0; Gaps 0;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1  CTGCTGCCCATCTGAATACAGAGATGGGGCTTGTGATTTCTCCACGGT 53
Db      155523  CTGCTGCCCATCTGAATACAGAGATGGGGCTTGTGATTTCTCCACGGT 155575

RESULT 15
LOCUS   AB090820/c
DEFINITION Anopheles gambiae retrotransposon R7Ag1 DNA, complete sequence.
ACCESSION AB090820
VERSION   AB090820.1 GI:28569877
KEYWORDS
SOURCE   Anopheles gambiae (African malaria mosquito)
ORGANISM
          Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
          Anopheles.
REFERENCE
  1  Kojima, K.K. and Fujiwara, H.
      Evolution of Target Specificity in R1 Clade Non-LTR
      Retrotransposons
      Mol. Biol. Evol. 20 (3), 351-361 (2003)
      12644555
  2  (bases 1 to 6470)
      Kojima, K.K. and Fujiwara, H.
      Direct Submission
      Submitted (26-AUG-2002) Kenji K Kojima, University of Tokyo,
      Department of Integrated Biosciences, Graduate School of Frontier
      Sciences; Bioscience Building 501, 5-1-5 Kashiwanoha, Kashiwa,
      Chiba 277-8562, Japan [E-mail:kk27513@mail.ecc.u-tokyo.ac.jp,
      Tel:81-4-7136-3661, Fax:81-4-7136-3660]
      Location/Qualifiers
        1..6470
          /organism="Anopheles gambiae"
          /mol_type="genomic DNA"
          /db_xref="taxon:7165"
        1..6470
          /note="non-LTR retrotransposon"
          /transposon="R7Ag1"
          959..2542
            /codon_start=1
            /product="gag-like protein"
            /protein_id="BAC57915.1"
            /db_xref="GI:28569878"
          /translation="MDQLRGRTISVDERPAVIRKLGSEKLGITVEEPSSAGVPAR
            TMAIGVKSAGTATKATSTPVSQGEVRMLAKADNETVIGVRKEEQIQLRLQ
            MEASNLQKEAREAREADARVREAREELREKELFNALLAOTLGGTSGARLES
            QOELQREQLLRMESQROEQQLEDDQQRQWRQOQQQQRQORLPAQOQWPTVOQS
            VRAQROGVTESSASVDEAGTWVVRGNQKQGNLQOQSAQROQPAHQOQW
            PQOQOQQOQQRNGIHQOQKRRPRKRPDEIVVPAPGVSKEMVVKIRTNRIADFO
            RQIGVGRTPRDHLLPLSRDVBDAALKDIIQEVIGERSVTVRTMAEVVLGIDNM
            IDEEAIKALMTTLGKQSLVATVNLWERDMTKARVLRPAEALVKDRRLGLGTY
            CSVHEAPKVSGLTRCFLCRGHIATCTGEDRSKRCRCRCGQDQTHRAGCTNEVKCM
            LCGGAHRIGAAACGGQPSN"
          2544..6212
            /note="ORF2"
            /codon_start=1
            /product="reverse transcriptase"
            /protein_id="BAC57916.1"
            /db_xref="GI:28569879"
          /translation="MEVLQINVRNRSQADLALNTMRVERADVCLMVELHVSPPRNNGN
            WVARDRGKVAIATASSEYVQVQSVTQSGIAAARINDVLFICCVYVSPAGVSEFEV
            MQRIDVLARGHPRVVFAGDLNAAWHTAWGSCRTNAGKEAVQLVDSLGLVNTGTAPT
            FLNGVAPRPSVDVAFASSLIAGVNTIPHRMRIISIYSNHHVIRFAVGELLQRP
            ADNRQEGSTRESCTRWTRHPDACLPGVALDVAFSTERTVSAESLERVMTACDAA
            MARVFPQSGHGRGPATWTPAIELNCRLAKERLEAAIDEEQIAAASDLIQVRTA
            LSSITTSKKEHFDLELAEDETCQWYRNLRLSGSWTAREDRSSVLEIVSTLF
            PQHPVDPASPQQLGEERPEVRDNEQLLDIASSLNPRKAPGLDGVNNAALTA
            IRKHTDIFKKLFOECLDNERFPDEWKQKALIPKPKPPGLASFRPILLNNPGKV
            YERLLLSRINDVLEDPRLAENQYGRGRGSTVQAIQLVVDAGSHAMSGFRNTNRD
            KRCLLVADVNRNFAFNTASWQCILATALEDKDGSTVFRQLNRILRDYFANRELIVYDTADGPV

```

```

TRRVTAGVPGQSILGPTLWNIMYDGLRVLPGLASVIGYADDIVVMARGCTPQEAAL
VAEQAVDAIAAMWEDHHLQLAPEKTEGVMISLRGQKVPFRVGTIIHQSQIRYL
GVIIHDHLSWKPHVELSTAKALRVVGVTVAVNRHSGPOVAKRRLIAAFAESIRYAA
PWSEATDLQWCKELAQVRPLARGVTSSFVSAVETGVAGLAGLVPRLLVREDARC
HRRLLAAPGASRKDIRLEERQGTFOEWQPAWDAAAAAPTASRYAWAHMIFDLHLWM
SRRHGEVDPHLSQVLTGHYFREYLHVCGFAPSAECPCPGSVESVAHVLFQCEVPH
LEVELSGYTSDFPNENNLIQMKLLESPERNWSIQAAARKITKVLQOLWREDEBLQNLQ
AHLAALPRAAAVDAGPLDGEQVSDGVAELFRSNRGRARRTRRRRAEERVEVRLA
SMAAAEREDSILMAAFAEEAGEAPPPIPMRRRGLPPSPRTVRAHERHRLYLQRL
YQORAREGTLPTVPHGRNRRSRAPSEADTIRRRMRREMERLRTARTARVPSNQGVRE
VLSADALAAITEATTSGR"

```

ORIGIN

Query Match 2.1%; Score 25; DB 3; Length 6470;
 Best Local Similarity 100.0%; Pred. No. 0.03;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 847 GCTGCTGCCGTGTGTTGTTTCTGCTG 871
 |||||
 Db 1574 GCTGCTGCCGTGTGTTGTTTCTGCTG 1550

Search completed: September 15, 2004, 03:47:20
 Job time : 4833.88 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 19:19:38 ; Search time 504.04 Seconds
(without alignments)
9903.249 Million cell updates/sec

Title: US-09-729-264-1

Perfect score: 1175

Sequence: 1 ctgtctgccatctgaataa.....gtaatacaactgtagtatag 1175

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5744508

Minimum DB seq length: 16

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- N_Geneseqn_29Jan04:*
- 1: Geneseqn1980s:*
 - 2: Geneseqn1990s:*
 - 3: Geneseqn2000s:*
 - 4: Geneseqn2001as:*
 - 5: Geneseqn2001bs:*
 - 6: Geneseqn2002s:*
 - 7: Geneseqn2003as:*
 - 8: Geneseqn2003bs:*
 - 9: Geneseqn2003cs:*
 - 10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1175	100.0	1175	6	ABK13028 DNA encod
2	1076	91.6	2051	9	ADB62950 Human cDN
3	1074	91.4	1240	6	ABK13030 DNA encod
4	1027	87.4	1168	6	ABK13029 DNA encod
5	739	62.9	1139	6	ABK13031 DNA encod
6	728	62.0	1392	5	AAS92356 DNA encod
7	217	18.5	474	8	ACH16130 Human adu
8	144	12.3	401	4	AAI36582 Probe #52
9	25	2.1	357	5	AAI04335 Probe #43
10	24	2.0	6507	4	ABL15493 Drosophil
11	24	2.0	59967	4	ABL15492 Drosophil
12	23	2.0	1195	6	ABK13032 DNA encod
13	22	1.9	141	2	AAV993359 5' PCR pr
14	22	1.9	767	6	ABN99162 Arabidops
15	22	1.9	1200	2	AAV35364 Human GDN
16	22	1.9	1200	2	AAV35365 Human GDN
17	22	1.9	1203	3	AAAI2547 DNA encod
18	22	1.9	1699	2	AAV00251 Human Ret
19	22	1.9	1792	3	AAZ29104 Human GPR
20	22	1.9	1809	2	AAV99333 Glial cel
21	22	1.9	1829	3	AZ93702 PRO538 DN
22	22	1.9	1829	3	AAZ98519 Human PRO
23	22	1.9	1829	3	AAZ77617 Human PRO

24	22	1.9	1829	3	AAZ54101 PRO538 cD
25	22	1.9	1829	3	AAC58234 Human PRO
26	22	1.9	1837	3	AAC229105 Human GFR
27	22	1.9	1837	3	AAZ88520 Human PRO
28	22	1.9	1837	3	AAC58235 Human PRO
29	22	1.9	1878	2	AAV99334 Glial cel
30	22	1.9	1991	2	AAV99329 cDNA enco
31	22	1.9	3942	2	AAX35851 Rat angio
32	22	1.9	4142	7	ABT41796 Toxicity
33	22	1.9	4142	9	ADB52524 Primary r
34	21	1.8	269	3	AAC10184 Human sec
35	21	1.8	493	8	ACH23839 Human adu
36	21	1.8	550	5	AAS05528 Mammalian
37	21	1.8	921	5	AAS80140 DNA encod
38	21	1.8	1126	6	ABQ61025 Human dif
39	21	1.8	1558	5	AAH81511 Human pol
40	21	1.8	1679	4	AAI93530 Human pol
41	21	1.8	1830	5	AAS80204 DNA encod
42	21	1.8	2649	5	AAS88493 Drosophil
43	21	1.8	2803	4	ABL10241 Drosophil
44	21	1.8	7782	4	ABL20110 Drosophil
45	21	1.8	8747	4	ABL10240 Drosophil

ALIGNMENTS

RESULT 1

ID ABK13028 standard; cDNA; 1175 BP.
XX
AC ABK13028;

DT 23-APR-2002 (first entry)

XX
DE DNA encoding human B7-like protein, B7-L_h1.
XX

Human; B7-like protein; B7-L; antiinfertility; gynaecological; antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic; antiinflammatory; dermatological; antipsoriatic; neuroprotective; antidiabetic; haemostatic; antithyroid; antitumor; antiallergic; antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer; reproductive disorder; graft versus host disease; autoimmune disease; toxic shock syndrome; allergy; nephropathy; skin disorder; endocrinopathy; lymphoproliferative disorder; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FH CDS 27..1175
FT /**tag= a
FT /product= "B7-like protein, B7-L_h1"

XX WO200200710-A2.

XX 03-JAN-2002.

XX 28-JUN-2001; 2001WO-US020719.

XX 28-JUN-2000; 2000US-0214512P.

XX 28-NOV-2000; 2000US-00729264.

XX (AMGE-) AMGEN INC.

XX Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;

XX WPI; 2002-130881/17.

XX P-PSDB; AAU75540.

XX New B7-like polypeptides, polynucleotides and their modulators, useful

XX for diagnosing, preventing and treating reproductive, immune and

XX proliferative disorders, e.g. cancer and arteriosclerosis.

PS	Claim 1; Fig 1; 135pp; English.	
XX	The invention relates to an isolated B7-like (B7-L) polypeptide (I). The	
CC	polypeptide, polynucleotide encoding it and antibody against (I) are	
CC	useful for treating B7-like polypeptide-related disease, disorders or	
CC	conditions including reproductive disorders (e.g. infertility,	
CC	miscarriage, preterm labour and delivery and endometriosis) and	
CC	proliferative disorders. Antibodies, soluble proteins comprising	
CC	extracellular domains and other regulators of B7-L polypeptides are	
CC	useful for enhancing the immune response to tumours. (I) plays a role in	
CC	growth and maintenance of cancer cells based on the observation of	
CC	seminal vesicle hyperplasia in transgenic mice overexpressing B7-L	
CC	polypeptide. Hence modulators of (I) are useful for the treatment of	
CC	cancer including seminal vesicle cancer, lung, brain, breast, ovarian,	
CC	testicular cancer and cancers of haematopoietic system. B7-L polypeptide	
CC	pathway can be manipulated to regulate cytotoxic T-lymphocyte response in	
CC	allograft transplantation, graft versus host disease, T-cell dependent B-	
CC	cell mediated diseases and autoimmune diseases. B7-L molecules are useful	
CC	for alleviating the symptoms associated with diseases involving chronic	
CC	immune cell dysfunction or to treat autoimmune diseases such as systemic	
CC	lupus erythematosus, rheumatoid arthritis, multiple sclerosis, diabetes,	
CC	immune thrombocytopenic purpura and psoriasis, chronic inflammatory	
CC	disease such as inflammatory bowel disease (Crohn's disease and	
CC	ulcerative colitis), Grave's disease, Hashimoto's thyroiditis and	
CC	diabetes mellitus. They are also useful as immunosuppressive agents for	
CC	bone marrow and organ transplantation or to prolong graft survival. B7-L	
CC	molecules are also useful for diagnosis and treatment of diseases	
CC	involving abnormal cell proliferation, including arteriosclerosis and	
CC	vascular restenosis. Antagonists of B7-L polypeptides are useful for	
CC	alleviation of toxic shock syndrome or all sensitisation due to blood	
CC	transfusions, and for treatment of allergy, asthma and hypersensitivity	
CC	reactions, nephropathies (e.g. glomerulonephritis), skin disorders	
CC	(pemphigus and pemphigoid), endocrinopathies (Grave's disease), various	
CC	pneumopathies (extrinsic alveolitis), vasculopathies, coeliac disease,	
CC	anaemia, thrombocytopenias, Guillain-Barre syndrome and myasthenia	
CC	gravis, and lymphoproliferative disorders such as multiple myeloma. The	
CC	present sequence represents the coding sequence of human B7-L _{h1}	
XX		
SQ	Sequence 1175 BP; 295 A; 307 C; 286 G; 287 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 1175; DB 6; Length 1175;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 1175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 CTGTCTGCCATCTGAATAAAGAGATGGGGCTTGTGATTTCTCCACGGTCTCGGT 60	
Db	1 CTGTCTGCCATCTGAATAAAGAGATGGGGCTTGTGATTTCTCCACGGTCTCGGT 60	
Qy	61 CTGTAATGAAGTCAATAGAGCCCGCCAGATGCAACAGTCTGAAGGGCTCCAGGCTC 120	
Db	61 CTGTAATGAAGTCAATAGAGCCCGCCAGATGCAACAGTCTGAAGGGCTCCAGGCTC 120	
Qy	121 GCTTCAACTGCACCGCTCTCCAGGGCTGGAAGCTCATCATGTTGGCTCTCAGTGACATGG 180	
Db	121 GCTTCAACTGCACCGCTCTCCAGGGCTGGAAGCTCATCATGTTGGCTCTCAGTGACATGG 180	
Qy	181 TGGTGCTAAGCGTCAAGCGCCATGGAGCCCATCATCAACCAATGACCGCTTCACTCTCAGA 240	
Db	181 TGGTGCTAAGCGTCAAGCGCCATGGAGCCCATCATCAACCAATGACCGCTTCACTCTCAGA 240	
Qy	241 GGTACGACAGGGCGGGAATTTCACTCGGATGATCATCATCAATGTGGAGCCAGTG 300	
Db	241 GGTACGACAGGGCGGGAATTTCACTCGGATGATCATCATCAATGTGGAGCCAGTG 300	
Qy	301 ATTCTGGGGAACATCAGATGAGCTCCAGAACAGTGGCTGATGATGATCTGCTTACCTTA 360	
Db	301 ATTCTGGGGAACATCAGATGAGCTCCAGAACAGTGGCTGATGATGATCTGCTTACCTTA 360	
Qy	361 CGGTCAAGTTATGGAGAGCTGTTCAATCCAGTGTAAATCTTGTAGTCGCTGGAATG 420	
Db	361 CGGTCAAGTTATGGAGAGCTGTTCAATCCAGTGTAAATCTTGTAGTCGCTGGAATG 420	
Qy	421 AACCTTGTGAAGTTACTTGTCTACCTCACACTGGACCCGGCTCCCGGATATTTCTGGG 480	

Db	421 AACCTTGTGAAGTTACTTGTCTACCTCACACTGGACCCGGCTCCCGGATATTTCTGGG 480	
Qy	481 AGCTCGGTCTCTGGTTCAGCCATTCAAGCTATTAATTTTCTCGGAGCCAGGACCTTC 540	
Db	481 AGCTCGGTCTCTGGTTCAGCCATTCAAGCTATTAATTTTCTCGGAGCCAGGACCTTC 540	
Qy	541 AAAGTGCAGTGAGCATCTCTGGTCTGACCCCAAGCAATGGGACTTTGACTTGGGTGG 600	
Db	541 AAAGTGCAGTGAGCATCTCTGGTCTGACCCCAAGCAATGGGACTTTGACTTGGGTGG 600	
Qy	601 CTACCTGGAAGAGCCCTGAAGGCCCGCAAGTCTGCAACTGTAAATCTCAGTGTATTCGGT 660	
Db	601 CTACCTGGAAGAGCCCTGAAGGCCCGCAAGTCTGCAACTGTAAATCTCAGTGTATTCGGT 660	
Qy	661 GTCCCCAAGACACTGAGAGTGGTATTAATATCCAGTGTATTAATCAAGTTTACCGAGTT 720	
Db	661 GTCCCCAAGACACTGAGAGTGGTATTAATATCCAGTGTATTAATCAAGTTTACCGAGTT 720	
Qy	721 TAGGTTTTCATTGCTTACTTTGGGCAAGTGGACTTGGACTAGCAGGACCATGCTTC 780	
Db	721 TAGGTTTTCATTGCTTACTTTGGGCAAGTGGACTTGGACTAGCAGGACCATGCTTC 780	
Qy	781 TGAGCGGAGCTGCTTACTTACAAATAGCTGCTGCTGCCGCCGCTGTTGTGTGGCT 840	
Db	781 TGAGCGGAGCTGCTTACTTACAAATAGCTGCTGCTGCCGCCGCTGTTGTGTGGCT 840	
Qy	841 GCAACTGCTCTGCGGTGTTGTTTCTGCTGTAGAGAAAAGAGGATTTTCGTTATCAAT 900	
Db	841 GCAACTGCTCTGCGGTGTTGTTTCTGCTGTAGAGAAAAGAGGATTTTCGTTATCAAT 900	
Qy	901 TTCAAAAAGAAATCTGAAAAAGAGAGCAACAAAGAAACTGAGACAGAAAGTGAATG 960	
Db	901 TTCAAAAAGAAATCTGAAAAAGAGAGCAACAAAGAAACTGAGACAGAAAGTGAATG 960	
Qy	961 AAAACTCGGCTACAATTCAGATCAAAAAGACCAAGACACAGCCGCTTCTCTCCCTCCCA 1020	
Db	961 AAAACTCGGCTACAATTCAGATCAAAAAGACCAAGACACAGCCGCTTCTCTCCCTCCCA 1020	
Qy	1021 AATCCTGTGAATCCAGTGATCTCTGAACAAAGAAACAGTAGCTGTGGCCCTCTCACCAGC 1080	
Db	1021 AATCCTGTGAATCCAGTGATCTCTGAACAAAGAAACAGTAGCTGTGGCCCTCTCACCAGC 1080	
Qy	1081 GGGCTGATCAACGTCACCCAGGCGAGCAAGTATCCACAGCTTCTTTTAACTCGGCCA 1140	
Db	1081 GGGCTGATCAACGTCACCCAGGCGAGCAAGTATCCACAGCTTCTTTTAACTCGGCCA 1140	
Qy	1141 GTCTGTGAAGGTCAGTAAATACAACTGTAGTATAG 1175	
Db	1141 GTCTGTGAAGGTCAGTAAATACAACTGTAGTATAG 1175	
RESULT 2		
ADB62950	ADB62950 standard; cDNA; 2051 BP.	
ID	ADB62950 standard; cDNA; 2051 BP.	
XX	XX	
AC	ADB62950;	
XX	XX	
DT	04-DEC-2003 (first entry)	
XX	XX	
DE	Human cDNA encoding clone PLACE60177880.	
XX	XX	
KW	Human; ss; gene; pharmaceutical; diagnostic; gene therapy;	
KW	tissue regeneration; cell regeneration; membrane protein;	
KW	signal transduction-related protein; transcription-related protein;	
KW	osteoporosis; neurological disease; cancer; tumour.	
XX	XX	
OS	Homo sapiens.	
XX	XX	
Key	Location/Qualifiers	
FT	104..1327	
FT	/tag= a	
FT	/product= "Clone PLACE60177880 protein"	

XX EN EP1308459-A2.
XX PD 07-MAY-2003.
XX PF 28-MAR-2002; 2002EP-00007401.
XX PR 05-NOV-2001; 2001JP-00379298.
XX PR 25-JAN-2002; 2002US-00350978.
XX PA (HELI-) HELIX RES INST.
XX PA (RELI-) RES ASSOC BIOTECHNOLOGY.
XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX WPI: 2003-450961/43.
XX DR P-PSDB; ADB64920.
XX PT New polynucleotides and polypeptides, useful for developing a diagnostic
XX PT marker or medicines for regulation of their expression and activity, or
XX PT as targets of gene therapy.
XX PS Claim 1; Page; 222pp; English.
XX CC The invention discloses a polynucleotide comprising a sequence selected
XX CC from 1970 fully defined nucleotide sequences which encode novel
XX CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
XX CC or its partial peptide, an antibody binding to the polypeptide or peptide
XX CC of the polynucleotide, immunologically assaying the polypeptide or
XX CC peptide of the polynucleotide by contacting the polypeptide or peptide
XX CC with the antibody of the encoded protein, and observing the binding
XX CC between the two, a transformant carrying the polynucleotide in an
XX CC expressible manner and an antisense polynucleotide. The oligonucleotide
XX CC is useful as a primer for synthesising the polynucleotide, or as a probe
XX CC for detecting the polynucleotide. The polynucleotides and encoded
XX CC proteins are useful as pharmaceutical agents and many disease-related
XX CC genes may be included in them, for developing a diagnostic marker or
XX CC medicines for regulation of their expression and activity, or as targets
XX CC of gene therapy. The genes are involved in tissue and/or cell
XX CC regeneration. Membrane proteins, signal transduction-related proteins,
XX CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
XX CC neurological diseases, cancer, tumours). The cDNA may be used to regulate
XX CC the activity or expression of the encoded protein to treat diseases. The
XX CC sequence presented is a cDNA of the invention. Note: Some of the sequence
XX CC data for this patent is not represented in the printed specification, but
XX CC is based on sequence information supplied by the European Patent Office.
SQ Sequence 2051 BP; 580 A; 463 C; 474 G; 534 T; 0 U; 0 Other;

Query Match 91.6%; Score 1076; DB 9; Length 2051;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1126; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 49 ACGGTTCTGGGTCTGGTAATGAAGTCATAGAAGGCCGCCAGAAATGCAACAGTCTCTGAAGG 108
DB 201 ACGGTTCTGGGTCTGGTAATGAAGTCATAGAAGGCCGCCAGAAATGCAACAGTCTCTGAAGG 260
QY 109 GCTCCAGGCTCCGCTTCACTGACCGTCTCCAGGCTGGAAGCTCATCATGTTGGGCTC 168
DB 261 GCTCCAGGCTCCGCTTCACTGACCGTCTCCAGGCTGGAAGCTCATCATGTTGGGCTC 320
QY 169 TCAGTGACATGGTGGTCTGAAGCGTCTAGGCGCCATGGAGCCCATCATCAATGACCGCT 228
DB 321 TCAGTGACATGGTGGTCTGAAGCGTCTAGGCGCCATGGAGCCCATCATCAATGACCGCT 380
QY 229 TCACCTCTCAGAGGTACGACCCAGGCGGGAACTTCACTCGGAGATGATCATCCACAATG 298
DB 381 TCACCTCTCAGAGGTACGACCCAGGCGGGAACTTCACTCGGAGATGATCATCCACAATG 440
QY 289 TGGAGCCGAGTGATTCGGGGAAACATCAGATGCGAGCCTCCAGAACAGTGGCGCTGCATGGAT 348

Db 441 TGGAGCCCGAGTATCGGGGAACATCAGATCGAGCCTCAGAACAGTCGCCCTGCGATGGAT 500
QY 349 CTGCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTCAITCCCAAGTGTAAATCTTTAG 408
Db 501 CTGCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTCAITCCCAAGTGTAAATCTTTAG 560
QY 409 TCGCTGGAATGAACCTTGTGAAGTACTTGTCTACCTCTACACTGGACCCGGCTCCCGG 468
Db 561 TCGCTGGAATGAACCTTGTGAAGTACTTGTCTACCTCTACACTGGACCCGGCTCCCGG 620
QY 469 ATATTTCTGGGAGCTCGGTCTCTCTGGTTCAGCCATTCAAGCTATTATTTTGTTCGGAGC 528
Db 621 ATATTTCTGGGAGCTCGGTCTCTCTGGTTCAGCCATTCAAGCTATTATTTTGTTCGGAGC 680
QY 529 CCAGGACCTTCAAAGTGCAGTGCATCTCTGGCTCTGACCCCAAGCAATGGGACTT 588
Db 681 CCAGGACCTTCAAAGTGCAGTGCATCTCTGGCTCTGACCCCAAGCAATGGGACTT 740
QY 589 TGACTTGGTGGCTACCTGGAGAGCCTGGAAGGCCCGCAAGTCTGCAACTGTAAATCTCA 648
Db 741 TGACTTGGTGGCTACCTGGAGAGCCTGGAAGGCCCGCAAGTCTGCAACTGTAAATCTCA 800
QY 649 CTGTGATTTCGGTGTCCCAAGACACTCGAGGTGCTATTATATTTCCAGGTGTATTATCAA 708
Db 801 CTGTGATTTCGGTGTCCCAAGACACTCGAGGTGCTATTATATTTCCAGGTGTATTATCAA 860
QY 709 GTTTACGAGTTAGTGTTCCTTTCATTGCTCTACTGGGCAAGTGTGGACTGGACTAGCAG 768
Db 861 GTTTACGAGTTAGTGTTCCTTTCATTGCTCTACTGGGCAAGTGTGGACTGGACTAGCAG 920
QY 769 GCACCATGCTCTGACCGCAGCTGTACTTCAATACGCTGCTGCTGCTGCTGCTGCTGCTG 828
Db 921 GCACCATGCTCTGACCGCAGCTGTACTTCAATACGCTGCTGCTGCTGCTGCTGCTGCTG 980
QY 829 GTTGTGCTGGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 888
Db 981 GTTGTGCTGGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1040
QY 889 TTCGTATTTCATTTTCAAAAGAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAG 948
Db 1041 TTCGTATTTCATTTTCAAAAGAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAG 1100
QY 949 AAAGTGAATGAACCTCGGCTTCAATTCAGATGAACAAAGACACACACACCGCTT 1008
Db 1101 AAAGTGAATGAACCTCGGCTTCAATTCAGATGAACAAAGACACACACACCGCTT 1160
QY 1009 CTCTCCCTCCCAATCTGTGAATCCAGTGCATCTGAAACAAAGAAACAGTAGCTGTGCGC 1068
Db 1161 CTCTCCCTCCCAATCTGTGAATCCAGTGCATCTGAAACAAAGAAACAGTAGCTGTGCGC 1220
QY 1069 CTCTCAGCAGCGGGTGTATCAACGCTCCACCCAGGCGCAGCAAGTATCCACAGGCTTCTT 1128
Db 1221 CTCTCAGCAGCGGGTGTATCAACGCTCCACCCAGGCGCAGCAAGTATCCACAGGCTTCTT 1280
QY 1129 TTAATCTGGCCAGTCTCGAAGGTCAGTAATCAACTGTAGTATAG 1175
Db 1281 TTAATCTGGCCAGTCTCGAAGGTCAGTAATCAACTGTAGTATAG 1327

RESULT 3
ID ABK13030
XX ABK13030 standard; cDNA; 1240 BP.
AC ABK13030;
XX
XX
DT 23-APR-2002 (first entry)
XX
DE DNA encoding human B7-like protein, B7-L_h3.
XX
XX Human; B7-like protein; B7-L; antiinfectility; gynaecological;
KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;
KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;

KW antidiabetic; haemostatic; antithyroid; antiulcer; antiallergic;
KW antiasthmatic; nephrotoxic; antibacterial; virucide; tumour; cancer;
KW reproductive disorder; graft versus host disease; autoimmune disease;
KW toxic shock syndrome; allergy; nephropathy; skin disorder;
KW endocrinopathy; lymphoproliferative disorder; gene; ss.
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
PH 80..1240
FT CDS /*tag= a
FT /product= "B7-like protein, B7-L_h3"
XX
XX WO200200710-A2.
XX
XX 03-JAN-2002.
XX
XX 28-JUN-2001; 2001WO-US020719.
XX
XX 28-JUN-2000; 2000US-0214512P.
PR 28-NOV-2000; 2000US-00729264.
XX
XX (AMGE-) AMGEN INC.
XX
XX Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
PI
XX WPI; 2002-130881/17.
DR P-PSDB; AAU75542.
XX
XX New B7-like polypeptides, polynucleotides and their modulators, useful
PT for diagnosing, preventing and treating reproductive, immune and
PT proliferative disorders, e.g. cancer and arteriosclerosis.
XX
XX Claim 1; Fig 3; 135pp; English.
XX
XX The invention relates to an isolated B7-like (B7-L) polypeptide (I). The
CC polypeptide, polynucleotide encoding it and antibody against (I) are
CC useful for treating B7-like polypeptide-related disease, disorders or
CC conditions including reproductive disorders (e.g. infertility, and
CC miscarriage, preterm labour and delivery and endometriosis) and
CC proliferative disorders. Antibodies, soluble proteins comprising
CC extracellular domains and other regulators of B7-L polypeptides are
CC useful for enhancing the immune response to tumours. (I) plays a role in
CC growth and maintenance of cancer cells based on the observation of
CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
CC polypeptide. Hence modulators of (I) are useful for the treatment of
CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response in
CC allograft transplantation, graft versus host disease, T-cell dependent B-
CC cell mediated diseases and autoimmune diseases. B7-L molecules are useful
CC for alleviating the symptoms associated with diseases involving chronic
CC immune cell dysfunction or to treat autoimmune diseases such as systemic
CC lupus erythematosus, rheumatoid arthritis, multiple sclerosis, diabetes,
CC immune thrombocytopenic purpura and psoriasis, chronic inflammatory
CC disease such as inflammatory bowel disease (Crohn's disease and
CC ulcerative colitis), Grave's disease, Hashimoto's thyroiditis and
CC diabetes mellitus. They are also useful as immunosuppressive agents for
CC bone marrow and organ transplantation or to prolong graft survival. B7-L
CC molecules are also useful for diagnosis and treatment of diseases
CC involving abnormal cell proliferation, including arteriosclerosis and
CC vascular restenosis. Antagonists of B7-L polypeptides are useful for
CC alleviation of toxic shock syndrome or allosensitisation due to blood
CC transfusions, and for treatment of allergy, asthma and hypersensitivity
CC reactions, nephropathies (e.g. glomerulonephritis), skin disorders
CC (pemphigus and pemphigoid), endocrinopathies (Grave's disease), various
CC pneumopathies (extrinsic alveolitis), vasculopathies, coeliac disease,
CC anaemia, thrombocytopenias, Guillain-Barre syndrome and myasthenia
CC gravis, and lymphoproliferative disorders such as multiple myeloma. The
XX present sequence represents the coding sequence of human B7-L_h3
XX
XX Sequence 1240 BP; 318 A; 319 C; 305 G; 298 T; 0 U; 0 Other;

Query Match		91.4%;	Score 1074;	DB 6;	Length 1240;
Best Local Similarity		99.9%;	Pred. No. 0;		
Matches 1124;		Conservative	0;	Mismatches	1; Indels 0; Gaps 0;
QY	51	GGTTCTGGGTCTGGTAATGAAGTCATAGAAGGCCGCCAGAACTGCAACAGTCTCTGAAGGGC	110		
DB	116	GGTTCTGGGTCTGGTAATGAAGTCATAGAAGGCCGCCAGAACTGCAACAGTCTCTGAAGGGC	175		
QY	111	TCCAGGCTCGCTTCAACTGCACCGCTCCAGGGGCTGGAAGCTCATCATGTGGGTCTC	170		
DB	176	TCCAGGCTCGCTTCAACTGCACCGCTCCAGGGGCTGGAAGCTCATCATGTGGGTCTC	235		
QY	171	AGTGACATGGTGGTGTCTAAGCGGTGAGGCCCATGAGGCCCATCATCAACCAATGACCGCTTC	230		
DB	236	AGTGACATGGTGGTGTCTAAGCGGTGAGGCCCATGAGGCCCATCATCAACCAATGACCGCTTC	295		
QY	231	ACCTCTCAGAGGTACGACAGCGGGGAACTTCACCTCGGAGATGATCATCCCAATGTG	290		
DB	296	ACCTCTCAGAGGTACGACAGCGGGGAACTTCACCTCGGAGATGATCATCCCAATGTG	355		
QY	291	GAGCCAGTGATTCCGGGGAACATCAGATGCAGCTCCAGAACAGTCCCTGCAATGATCT	350		
DB	356	GAGCCAGTGATTCCGGGGAACATCAGATGCAGCTCCAGAACAGTCCCTGCAATGATCT	415		
QY	351	GCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCAATCCCAAGTGTAACTTGTAGTC	410		
DB	416	GCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCAATCCCAAGTGTAACTTGTAGTC	475		
QY	411	GCTGAGATGAACCTTGTGAAGTTACTGTCTACCTTCACACTGGACCCGGCTCCCGGAT	470		
DB	476	GCTGAGATGAACCTTGTGAAGTTACTGTCTACCTTCACACTGGACCCGGCTCCCGGAT	535		
QY	471	ATTTCTCTGGGAGCTCGGTCTCTCGGTGAGCCATTCAAAGCTATTATTTTGTCCGAGCCC	530		
DB	536	ATTTCTCTGGGAGCTCGGTCTCTCGGTGAGCCATTCAAAGCTATTATTTTGTCCGAGCCC	595		
QY	531	AGCGACCTTCAAAGTCAGTGAGCATCTCGCTCTGACCCACAGACAAATGGGACTTG	590		
DB	596	AGCGACCTTCAAAGTCAGTGAGCATCTCGCTCTGACCCACAGACAAATGGGACTTG	655		
QY	591	ACTTGGTGGGTACTCTGGAAGAGCTCAAGCCCGCAAGTCTGCAACTGTAATCTCACT	650		
DB	656	ACTTGGTGGGTACTCTGGAAGAGCTCAAGCCCGCAAGTCTGCAACTGTAATCTCACT	715		
QY	651	GTGATTCGGTGTCCCCAAGACACTGGAGGTGGTATTAAATATTCCAGTGTATTATCAAGT	710		
DB	716	GTGATTCGGTGTCCCCAAGACACTGGAGGTGGTATTAAATATTCCAGTGTATTATCAAGT	775		
QY	711	TTACCGAGTTTGTAGTTTTCATTCCTACTCTGGGGCAAGTTGGACTTGGACTAGCAGGC	770		
DB	776	TTACCGAGTTTGTAGTTTTCATTCCTACTCTGGGGCAAGTTGGACTTGGACTAGCAGGC	835		
QY	771	ACCATGCTCTGACCGCAGCTGTACTCTTACAATACGCTGCTGCTGCGCGCGTCTGT	830		
DB	836	ACCATGCTCTGACCGCAGCTGTACTCTTACAATACGCTGCTGCTGCGCGCGTCTGT	895		
QY	831	TGTTGTGGTGCATCTGCTGCTGCTGTTGTTGTTCTGCTGTAGAGAAAAGAGGATTT	890		
DB	896	TGTTGTGGTGCATCTGCTGCTGCTGTTGTTGTTCTGCTGTAGAGAAAAGAGGATTT	955		
QY	891	CGTATTCAATTTCAAAGAAATCTGMAAAGAGAGAGCAACAAGAACTGAGACAGAA	950		
DB	956	CGTATTCAATTTCAAAGAAATCTGMAAAGAGAGAGCAACAAGAACTGAGACAGAA	1015		
QY	951	AGTGAATGAATACTCCGGCTACAATTCAGATGAACAAAAGAGACACAGACACCGCTTCT	1010		
DB	1016	AGTGAATGAATACTCCGGCTACAATTCAGATGAACAAAAGAGACACAGACACCGCTTCT	1075		
QY	1011	CTCCCTCCCAATCTCTGTAATCCAGTGAATCTCTGAAACAAAGAACAGTAGCTGTGCCCT	1070		
DB	1076	CTCCCTCCCAATCTCTGTAATCCAGTGAATCTCTGAAACAAAGAACAGTAGCTGTGCCCT	1135		
QY	1071	CCTCACAGCGGGCTGATCAACGTCACCCAGGCCAGCAAGTCTCCACAGGCTTCTTTT	1130		

Db 1136 CCTCACCAGGGGCTGATCAACGTCACCCAGCCAGCAAGTCATCCACAGGCTTCTTTT 1195
QY 1131 AATCTGCCAGTCTCTGAGAGGTCAGTAATACAACTGTAGTATAG 1175
Db 1196 AATCTGCCAGTCTCTGAGAGGTCAGTAATACAACTGTAGTATAG 1240
RESULT 4
ABK13029
ID ABK13029 standard; cDNA; 1168 BP.
AC ABK13029;
XX
DT 23-APR-2002 (first entry)
DE
DE DNA encoding human B7-like protein, B7-L_h2.
XX
KW Human; B7-like protein; B7-L; antiinfertility; gynaecological;
KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;
KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;
KW antidiabetic; haemostatic; antithyroid; antiulcer; antiallergic;
KW antiashtatic; nephrotropic; antibacterial; virucide; tumour; cancer;
KW reproductive disorder; graft versus host disease; autoimmune disease;
KW toxic shock syndrome; allergy; nephropathy; skin disorder;
KW endocrinopathy; lymphoproliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
FH CDS 8..1168
FT /*tag= a
FT /product= "B7-like protein, B7-L_h2"
FT
FT
XX
XX WO200200710-A2.
XX
XX
XX
XX 03-JAN-2002.
XX
XX
XX 28-JUN-2001; 2001WO-US020719.
XX
XX 28-JUN-2000; 2000US-0214512P.
XX 28-NOV-2000; 2000US-00729264.
XX
XX (AMGE-) AMGEN INC.
XX
XX
XX Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
XX
XX WPI; 2002-130881/17.
XX P-PSDB; AAU75541.
XX
XX New B7-like polypeptides, polynucleotides and their modulators, useful
XX for diagnosing, preventing and treating reproductive, immune and
XX proliferative disorders, e.g. cancer and arteriosclerosis.
XX
XX Claim 1; Fig 2; 135pp; English.
XX
XX
XX The invention relates to an isolated B7-like (B7-L) polypeptide (I). The
XX polypeptide, polynucleotide encoding it and antibody against (I) are
XX useful for treating B7-like polypeptide-related disease, disorders or
XX conditions including reproductive disorders (e.g. infertility,
XX miscarriage, preterm labour and delivery and endometriosis) and
XX proliferative disorders. Antibodies, soluble proteins comprising
XX extracellular domains and other regulators of B7-L polypeptides are
XX useful for enhancing the immune response to tumours. (I) plays a role in
XX growth and maintenance of cancer cells based on the observation of
XX seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
XX polypeptide. Hence modulators of (I) are useful for the treatment of
XX cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
XX testicular cancer and cancers of haematopoietic system. B7-L polypeptide
XX pathway can be manipulated to regulate cytotoxic T-lymphocyte response in
XX allograft transplantation, graft versus host disease, T-cell dependent B-
XX cell mediated diseases and autoimmune diseases. B7-L molecules are useful
XX for alleviating the symptoms associated with diseases involving chronic

CC immune cell dysfunction or to treat autoimmune diseases such as systemic
CC lupus erythematosus, rheumatoid arthritis, multiple sclerosis, diabetes,
CC immune thrombocytopenic purpura and psoriasis, chronic inflammatory
CC disease such as inflammatory bowel disease (Crohn's disease and
CC ulcerative colitis), Grave's disease, Hashimoto's thyroiditis and
CC diabetes mellitus. They are also useful as immunosuppressive agents for
CC bone marrow and organ transplantation or to prolong graft survival. B7-L
CC molecules are also useful for diagnosis and treatment of diseases
CC involving abnormal cell proliferation, including arteriosclerosis and
CC vascular restenosis. Antagonists of B7-L polypeptides are useful for
CC alleviation of toxic shock syndrome or allo sensitisation due to blood
CC transfusions, and for treatment of allergy, asthma and hypersensitivity
CC reactions, nephropathies (e.g. glomerulonephritis), skin disorders
CC (pemphigus and pemphigoid), endocrinopathies (Grave's disease), various
CC pneumopathies (extrinsic alveolitis), vasculopathies, coeliac disease,
CC anaemia, thrombocytopenias, Guillain-Barre syndrome and myasthenia
CC gravis, and lymphoproliferative disorders such as multiple myeloma. The
CC present sequence represents the coding sequence of human B7-L_h2
XX

SQ Sequence 1168 BP; 299 A; 302 C; 288 G; 279 T; 0 U; 0 Other;
Query Match 87.4%; Score 1027; DB 6; Length 1168;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1077; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGTCCTGAGGGCTCCAGGCTCGCTTCAACGTGACCGTCTCCAGGGCTGGAAGCTCAT 157
Db 91 AGTCCTGAGGGCTCCAGGCTCGCTTCAACGTGACCGTCTCCAGGGCTGGAAGCTCAT 150
QY 158 CATGTGGGCTCTCAGTGACATGGTGTCTAAGCGTCAGGCGCCATGGAGCCCATCATCAC 217
Db 151 CATGTGGGCTCTCAGTGACATGGTGTCTAAGCGTCAGGCGCCATGGAGCCCATCATCAC 210
QY 218 CAATGACCGCTTCACTCTCAGAGGTACGACAGGGCGGGAACCTTCCTCGGAGATGAT 277
Db 211 CAATGACCGCTTCACTCTCAGAGGTACGACAGGGCGGGAACCTTCCTCGGAGATGAT 270
QY 278 CATCCACAATGTGAGGCCAGTGATTCGGGGAACATCAGATGCGCCTCCAGAACAGTCG 337
Db 271 CATCCACAATGTGAGGCCAGTGATTCGGGGAACATCAGATGCGCCTCCAGAACAGTCG 330
QY 338 CCTGCATGGATCTGCTTACCTTACCGTCCAGTTCAGGAGCTGTTTCATTCAGTGT 397
Db 331 CCTGCATGGATCTGCTTACCTTACCGTCCAGTTCAGGAGCTGTTTCATTCAGTGT 390
QY 398 TAATCTTTGTAGTCGCTGAGAAATGAACCTTGTGAAGTTACTTGTACCTCACAACCTGGAC 457
Db 391 TAATCTTTGTAGTCGCTGAGAAATGAACCTTGTGAAGTTACTTGTACCTCACAACCTGGAC 450
QY 458 CCGGCTCCCGGATATTCCTGGGAGCTCGGTCTCCTGGTCAGCCATTCAAGCTATTATTT 517
Db 451 CTGGCTCCCGGATATTCCTGGGAGCTCGGTCTCCTGGTCAGCCATTCAAGCTATTATTT 510
QY 518 TGTTCGGAGCCAGCGACCTTCAAAAGTCAGTGAGCATCTGGCTCTGACCCACAGAG 577
Db 511 TGTTCGGAGCCAGCGACCTTCAAAAGTCAGTGAGCATCTGGCTCTGACCCACAGAG 570
QY 578 CAATGGGACTTTCACCTTGGCTGCTACCTGGAAGAGCCTGAAGCGCCGCAAGTCTGCAAC 637
Db 571 CAATGGGACTTTCACCTTGGCTGCTACCTGGAAGAGCCTGAAGCGCCGCAAGTCTGCAAC 630
QY 638 TGTAAATCTCAGTGTGATTCGGTGTCCCAAGACATCGGAGGTGGTATTAAATTCAGG 697
Db 631 TGTAAATCTCAGTGTGATTCGGTGTCCCAAGACATCGGAGGTGGTATTAAATTCAGG 690
QY 698 TGTATTATCAAGTTTACCAGTTTACGTTTTCATTGCTCTTGGGCAAGTGGACT 757
Db 691 TGTATTATCAAGTTTACCAGTTTACGTTTTCATTGCTCTTGGGCAAGTGGACT 750
QY 758 TGGACTAGCAGGACCATGCTTCTGACGCGGAGCTGTACTTCTTACAAATACGCTGCTG 817
Db 751 TGGACTAGCAGGACCATGCTTCTGACGCGGAGCTGTACTTCTTACAAATACGCTGCTG 810

QY 818 CTGCGCGCGTCTGTTGTTGCTGCAACTGCTGCTGCGCTGTTGTTCTGCTGTAGAG 877
DB |||||||
QY 811 CTGCGCGCGTCTGTTGTTGCTGCAACTGCTGCTGCGCTGTTGTTCTGCTGTAGAG 870
DB |||||||
QY 878 AAAAGAGGATTCGTATTCAATTTCAAAGAAATCTGAAAAGAGAAACAAAGA 937
DB |||||||
QY 871 AAAAGAGGATTCGTATTCAATTTCAAAGAAATCTGAAAAGAGAAACAAAGA 930
DB |||||||
QY 938 AACTGAGACAGAAAGTGGAAATGAAATCTCGGCTGAGATTGAGATGAACAAAGCCAC 997
DB |||||||
QY 931 AACTGAGACAGAAAGTGGAAATGAAATCTCGGCTGAGATTGAGATGAACAAAGCCAC 990
DB |||||||
QY 998 AGACACGCTTCTCTCCCTCCCAATCTCTGAAATCAGTGATCTCTGAAACAAAGAAACAG 1057
DB |||||||
QY 991 AGACACGCTTCTCTCCCTCCCAATCTCTGAAATCAGTGATCTCTGAAACAAAGAAACAG 1050
DB |||||||
QY 1058 TAGCTGTGGCCCTCTCCACAGCGGGCTGATCAACGCTCCACCGGCGCAAGTCAATCC 1117
DB |||||||
QY 1051 TAGCTGTGGCCCTCTCCACAGCGGGCTGATCAACGCTCCACCGGCGCAAGTCAATCC 1110
DB |||||||
QY 1118 ACAGGCTTCTTTAAATCTGGCCAGTCTGAGAGGTCAGTAATACAACTGATGATAG 1175
DB |||||||
QY 1111 ACAGGCTTCTTTAAATCTGGCCAGTCTGAGAGGTCAGTAATACAACTGATGATAG 1168
DB |||||||

RESULT 5

ABK13031

ID ABK13031 standard; cDNA; 1139 BP.

AC ABK13031;

XX 23-APR-2002 (first entry)

DT

DE

XX DNA encoding human B7-like protein, B7-L_{h4}.

Human; B7-like protein; B7-L; antiinfertility; gynaecological;
antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;
antiinflammatory; dermatological; antipsoriatic; neuroprotective;
antidiabetic; haemostatic; antithyroid; antiulcer; antiallergic;
antiasthmatic; nephrotoxic; antibacterial; virucide; tumour; cancer;
reproductive disorder; graft versus host disease; autoimmune disease;
toxic shock syndrome; allergy; nephropathy; skin disorder;
endocrinopathy; lymphoproliferative disorder; gene; ss.

XX Homo sapiens.

XX Key

XX Location/Qualifiers

FT 1. 1134

FT /*tag= a

FT /product= "B7-like protein, B7-L_{h4}"

XX

XX WO200200710-A2.

XX

XX 03-JAN-2002.

XX 28-JUN-2001; 2001WO-US020719.

XX 28-JUN-2000; 2000US-0214512P.

XX 28-NOV-2000; 2000US-00729264.

XX (AMGE-) AMGEN INC.

XX

XX Welcher AA, Sarmiento UM, Schultz HU, Chute HT;

XX WPI; 2002-130881/17.

XX New B7-like polypeptides, polynucleotides and their modulators, useful

XX for diagnosing, preventing and treating reproductive, immune and

XX proliferative disorders, e.g. cancer and arteriosclerosis.

XX Claim 1; Fig 4; 135pp; English.

XX The invention relates to an isolated B7-like (B7-L) polypeptide (I). The

CC polypeptide, polynucleotide encoding it and antibody against (I) are
CC useful for treating B7-like polypeptide-related disease, disorders or
CC conditions including reproductive disorders (e.g. infertility,
CC miscarriage, preterm labour and delivery and endometriosis) and
CC proliferative disorders. Antibodies, soluble proteins comprising
CC extracellular domains and other regulators of B7-L polypeptides are
CC useful for enhancing the immune response to tumours. (I) plays a role in
CC growth and maintenance of cancer cells based on the observation of
CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
CC cancer, including seminal vesicle cancer, lung, brain, breast, ovarian,
CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response in
CC allograft transplantation, graft versus host disease, T-cell dependent B-
CC cell mediated diseases and autoimmune diseases. B7-L molecules are useful
CC for alleviating the symptoms associated with diseases involving chronic
CC immune cell dysfunction or to treat autoimmune diseases such as systemic
CC lupus erythematosus, rheumatoid arthritis, multiple sclerosis, diabetes,
CC immune thrombocytopenic purpura and psoriasis, chronic inflammatory
CC disease such as inflammatory bowel disease (Crohn's disease and
CC ulcerative colitis), Grave's disease, Hashimoto's thyroiditis and
CC diabetes mellitus). They are also useful as immunosuppressive agents for
CC bone marrow and organ transplantation or to prolong graft survival. B7-L
CC molecules are also useful for diagnosis and treatment of diseases
CC involving abnormal cell proliferation, including arteriosclerosis and
CC vascular restenosis. Antagonists of B7-L polypeptides are useful for
CC alleviation of toxic shock syndrome or allo sensitisation due to blood
CC transfusions, and for treatment of allergy, asthma and hypersensitivity
CC reactions, nephropathies (e.g. glomerulonephritis), skin disorders
CC (pemphigus and pemphigoid), endocrinopathies (Grave's disease), various
CC pneumopathies (extrinsic alveolitis), vasculopathies, coeliac disease,
CC anaemia, thrombocytopenias, Guillain-Barre syndrome and myasthenia
CC gravis, and lymphoproliferative disorders such as multiple myeloma. The
CC present sequence represents the coding sequence of human B7-L_{h4}
XX
SQ

Sequence 1139 BP; 290 A; 300 C; 283 G; 266 T; 0 U; 0 Other;

Query Match 62.9%; Score 739; DB 6; Length 1139;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 789; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGTCCTGAAGGGCTCCAGGGCTCGCTTCAACTGACACCGTCTCCAGGGCTGGAAGCTCAT 157

DB 84 AGTCCTGAAGGGCTCCAGGGCTCGCTTCAACTGACACCGTCTCCAGGGCTGGAAGCTCAT 143

QY 158 CATGTGGGCTCTCAGTGACATGTTGGTCTAAGCGTTCAGGCCCTATGGAGCCCATCATCAC 217

DB 144 CATGTGGGCTCTCAGTGACATGTTGGTCTAAGCGTTCAGGCCCTATGGAGCCCATCATCAC 203

QY 218 CAATGACCGCTTACCTCTCAGAGGTACGACCGGGGAACTTCACTCGGAGATGAT 277

DB 204 CAATGACCGCTTACCTCTCAGAGGTACGACCGGGGAACTTCACTCGGAGATGAT 263

QY 278 CATCCACAATGTGGAGCCAGTGTTCGGGGAAATCAGATGCGACCTCCAGAACAGTCG 337

DB 264 CATCCACAATGTGGAGCCAGTGTTCGGGGAAATCAGATGCGACCTCCAGAACAGTCG 323

QY 338 CCTGATGGATCTGCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTCATCCAGTGT 397

DB 324 CCTGATGGATCTGCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTCATCCAGTGT 383

QY 398 TAACTTTGTAGTCGCTGAGATGAACCTTGTGAAGTTACTTGTCTACCTCATCTGGAC 457

DB 384 TAACTTTGTAGTCGCTGAGATGAACCTTGTGAAGTTACTTGTCTACCTCATCTGGAC 443

QY 458 CCGGCTCCCGGATATTTCTCTGGAGCTCGGTCTCTCTGTGTGAGCCATTCAAGCTATTATTT 517

DB 444 CCGGCTCCCGGATATTTCTCTGGAGCTCGGTCTCTCTGTGTGAGCCATTCAAGCTATTATTT 503

QY 518 TGTTCGGAGCCAGCGACCTTCAAAGTGCAGTGAGCTCTGGGCTCTGACCCACAGAG 577

DB 504 TGTTCGGAGCCAGCGACCTTCAAAGTGCAGTGAGCTCTGGGCTCTGACCCACAGAG 563

QY 578 CAATGGGACTTTGACTTGGTGGCTACCTGGAAGAGCCCTGAAGCCCGCAAGTCTGCAAC 637
Db 564 CAATGGGACTTTGACTTGGTGGCTACCTGGAAGAGCCCTGAAGCCCGCAAGTCTGCAAC 623
QY 638 TGTAAATCTCACTGTGATTCGGTGTCCCAAGACACTGAGGTGGTATTAATATCCAGG 697
Db 624 TGTAAATCTCACTGTGATTCGGTGTCCCAAGACACTGAGGTGGTATTAATATCCAGG 683
QY 698 TGTATATCAAGTTTACCGAGTTTAGTGTTCATTCGCTACTTGGGGCAAAAGTTGGACT 757
Db 684 TGTATATCAAGTTTACCGAGTTTAGTGTTCATTCGCTACTTGGGGCAAAAGTTGGACT 743
QY 758 TGGACTAGCAGCACCATGCTTCTGACGCGGACGCTGACTTCAATACAGCTGCTGCTG 817
Db 744 TGGACTAGCAGCACCATGCTTCTGACGCGGACGCTGACTTCAATACAGCTGCTGCTG 803
QY 818 CTGCGCGCGTCTGCTGTGGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 877
Db 804 CTGCGCGCGTCTGCTGTGGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 863
QY 878 AAAAAGAGGA 887
Db 864 AAAAAGAGGA 873

RESULT 6

AAS92356
ID AAS92356 standard; cDNA; 1392 BP.

XX AC AAS92356;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #28160.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.

XX FN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABG28169.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 1; SEQ ID NO 28160; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridization probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1392 BP; 338 A; 361 C; 341 G; 352 T; 0 U; 0 Other;

Query Match 62.0%; Score 728; DB 5; Length 1392;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 ATGTGGGCTCTCAGTGACATGGTGGTGTCTAAGCGTCAGGCCCATGAGCCCATCATCACC 218
Db 1 ATGTGGGCTCTCAGTGACATGGTGGTGTCTAAGCGTCAGGCCCATGAGCCCATCATCACC 60

QY 219 AATGACCGCTTCACTCTCAGAGGTACGACAGGGCGGAACTTCACCTCGAGATGATC 278
Db 61 AATGACCGCTTCACTCTCAGAGGTACGACAGGGCGGAACTTCACCTCGAGATGATC 120

QY 279 ATCCAAATGTGGAGCCCACTGATTGGGGAACATCAGATCAGCCTCCAGAACAGTGGC 338
Db 121 ATCCAAATGTGGAGCCCACTGATTGGGGAACATCAGATCAGCCTCCAGAACAGTGGC 180

QY 339 CTGATGATCTGCTTACCTTACCGTCCAACTTATGGAGAGCTGTTCAATCCAGTGT 398
Db 181 CTGATGATCTGCTTACCTTACCGTCCAACTTATGGAGAGCTGTTCAATCCAGTGT 240

QY 399 AATCTGTAGTCGCTGAGATGAACCTTGTCAAGTTACTTCTACCTCCACACTGGACC 458
Db 241 AATCTGTAGTCGCTGAGATGAACCTTGTCAAGTTACTTCTACCTCCACACTGGACC 300

QY 459 CGGCTCCGGATATTTCTTGGAGCTCGGTCTCTCTGTCAGCCATTCAAGCTATTTT 518
Db 301 CGGCTCCGGATATTTCTTGGAGCTCGGTCTCTCTGTCAGCCATTCAAGCTATTTT 360

QY 519 GTTCCGGAGCCCGAGCCTTCAAGTGCAGTGAGCATCTGGCTCTGACCCACAGAGC 578
Db 361 GTTCCGGAGCCCGAGCCTTCAAGTGCAGTGAGCATCTGGCTCTGACCCACAGAGC 420

QY 579 AATGGGACTTTGACTTGGTGGCTTACCTGGAAGAGCCTGAAGGCCCGCAAGTCTGCAACT 638
Db 421 AATGGGACTTTGACTTGGTGGCTTACCTGGAAGAGCCTGAAGGCCCGCAAGTCTGCAACT 480

QY 639 GTAAATCTCACTGTGATTCGGTGTCCCAAGACACACTGGAGTGGTATTAATATTCAGGT 698
Db 481 GTAAATCTCACTGTGATTCGGTGTCCCAAGACACACTGGAGTGGTATTAATATTCAGGT 540

QY 699 GTATTATCAAGTTTACCGAGTTTAGTGTTCATTTGCTACTTGGGGCAAAAGTTGGACT 758
Db 541 GTATTATCAAGTTTACCGAGTTTAGTGTTCATTTGCTACTTGGGGCAAAAGTTGGACT 600

QY 759 GGACTAGCAGGCACCATGCTTCTGACGCCGACCTGACTCTTACAATACGCTGCTGCTGC 818
Db 601 GGACTAGCAGGCACCATGCTTCTGACGCCGACCTGACTCTTACAATACGCTGCTGCTGC 660

QY 819 TGGCGCGCTGTTGTGTGGCTGCAACTGCTGCTGCGTGTGTTGTTCTGCTAGAGA 878
Db 661 TGGCGCGCTGTTGTGTGGCTGCAACTGCTGCTGCGTGTGTTGTTCTGCTAGAGA 720

QY 879 AAAAGAGG 886
Db 721 AAAAGAGG 728

RESULT 7
ACH16130

XX	Drosophila melanogaster expressed polynucleotide SEQ ID NO 40961.
DE	Drosophila; developmental biology; cell signalling; insecticide;
XX	pharmaceutical; gene; ss.
KW	Drosophila melanogaster.
XX	Drosophila melanogaster.
OS	WO200171042-A2.
XX	27-SEP-2001.
XX	23-MAR-2001; 2001WO-US009231.
PF	23-MAR-2000; 2000US-0191637P.
XX	11-JUL-2000; 2000US-00614150.
PR	(PEKE) PE CORP NY.
XX	Venter JC, Adams M, Li PWD, Myers EW;
PI	WPI; 2001-656860/75.
XX	P-PSDB; ABB711390.
DR	New isolated nucleic acid detection reagent for detecting 1000 or more
XX	genes from Drosophila and for elucidating cell signaling and cell-cell
PT	interactions.
PT	Claim 1; SEQ ID NO 40961; 2lpp + Sequence Listing; English.
PS	The invention relates to an isolated nucleic acid detection reagent
XX	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA
CC	sequences (ABLI01840-ABLI16175) and the encoded proteins (ABB57737-
CC	ABB72072). The sequence data for this patent did not form part of the
CC	printed specification, but was obtained in electronic format directly
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	Sequence 6507 BP; 1745 A; 1888 C; 1712 G; 1162 T; 0 U; 0 Other;
XX	Query Match 2.0%; Score 24; DB 4; Length 6507;
XX	Best Local Similarity 100.0%; Pred. No. 0.26;
XX	Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Oy	807 CGCTGCTGCTGCTGCCGCCGTCGT 830
Db	
Db	2946 CGCTGCTGCTGCTGCCGCCGTCGT 2923
RESULT 11	
ABLI5492/c	
ID	ABLI5492 standard; cDNA; 59967 BP.
XX	ABLI5492;
AC	
XX	26-MAR-2002 (first entry)
DT	
XX	Drosophila melanogaster expressed polynucleotide SEQ ID NO 40958.
DE	Drosophila; developmental biology; cell signalling; insecticide;
XX	pharmaceutical; gene; ss.
KW	Drosophila melanogaster.
XX	WO200171042-A2.
XX	27-SEP-2001.
XX	23-MAR-2001; 2001WO-US009231.
PF	
XX	

(AMGE-) AMGEN INC.

Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
WPI; 2002-130881/17.
P-PSDB; AAU75544.

New B7-like polypeptides, polynucleotides and their modulators, useful for diagnosing, preventing and treating reproductive, immune and proliferative disorders, e.g. cancer and arteriosclerosis.

Claim 1; Fig 5; 135pp; English.

The invention relates to an isolated B7-like (B7-L) polypeptide (I). The polypeptide, polynucleotide encoding it and antibody against (I) are useful for treating B7-like polypeptide-related disease, disorders or conditions including reproductive disorders (e.g. infertility, miscarriage, preterm labour and delivery and endometriosis) and proliferative disorders. Antibodies, soluble proteins comprising extracellular domains and other regulators of B7-L polypeptides are useful for enhancing the immune response to tumours. (I) plays a role in growth and maintenance of cancer cells based on the observation of seminal vesicle hyperplasia in transgenic mice overexpressing B7-L polypeptide. Hence modulators of (I) are useful for the treatment of cancer including seminal vesicle cancer, lung, brain, breast, ovarian, testicular cancer and cancers of haematopoietic system. B7-L polypeptide pathway can be manipulated to regulate cytotoxic T-lymphocyte response in allograft transplantation, graft versus host disease, T-cell dependent B-cell mediated diseases and autoimmune diseases. B7-L molecules are useful for alleviating the symptoms associated with diseases involving chronic immune cell dysfunction or to treat autoimmune diseases such as systemic lupus erythematosus, rheumatoid arthritis, multiple sclerosis, diabetes, immune thrombocytopenic purpura and psoriasis, chronic inflammatory disease such as inflammatory bowel disease (Crohn's disease and ulcerative colitis), Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They are also useful as immunosuppressive agents for bone marrow and organ transplantation or to prolong graft survival. B7-L molecules are also useful for diagnosis, including arteriosclerosis and involving abnormal cell proliferation, antagonists of B7-L polypeptides are useful for alleviation of toxic shock syndrome or allosensitisation due to blood transfusions, and for treatment of allergy, asthma and hypersensitivity reactions, nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and pemphigoid), endocrinopathies (Grave's disease), various pneumopathies (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia, thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and lymphoproliferative disorders such as multiple myeloma. The present sequence represents the coding sequence of mouse B7-L_{ml}

Sequence 1195 BP; 313 A; 304 C; 291 G; 287 T; 0 U; 0 Other;

Query Match 2.0%; Score 23; DB 6; Length 1195;
Best Local Similarity 100.0%; Pred. NO. 0.81;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 462 CTCCTCGGATATTTCCTGGGAGCT 484
|||||
DB 521 CTCCTCGGATATTTCCTGGGAGCT 543
|||||

RESULT 13
AAV99359
ID ID AAV99359 standard; DNA; 141 BP.
XX AC AAV99359;
XX AC
XX AC
DT DT
DE 25-MAR-1999 (first entry)
XX 5' PCR primer for GDNFR-gammal cDNA encoding the full length protein.
XX Glial cell line-derived neurotrophic factor receptor gamma 1;
KW GDNFR-alpha; glial cell line-derived neurotrophic factor; GDNF;
KW neurodegenerative disease; amyotrophic lateral sclerosis; GDNFR-gammal;

CC that may affect the biological function of the gene or gene products.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=999909770445

XX
SQ Sequence 767 BP; 169 A; 199 C; 147 G; 252 T; 0 U; 0 Other;

Query Match 1.9%; Score 22; DB 6; Length 767;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 821 CGCGCTGCTGTTGTGGCTGC 842
Db 672 CGCGCTGCTGTTGTGGCTGC 693

RESULT 15

AAV5364
ID AAV5364 standard; cDNA; 1200 BP.

XX
AC AAV5364;

XX
DT 28-SEP-1998 (first entry)

XX
DE Human GDNF alpha-3 receptor cDNA #1.

XX
KW Glial cell line-derived neurotrophic factor alpha-3 receptor; GDNF;
KW treatment; neurodegenerative disease; Parkinson's Disease; ALS; SMA;
KW amyotrophic lateral sclerosis; spinal muscular atrophy; nerve; trauma;
KW Huntington's Disease; Alzheimer's Disease; diabetic neuropathy; muscle;
KW muscular dystrophy; diagnostic; ss.

XX
OS Homo sapiens.

XX
FH Key Location/Qualifiers
FT CDS 1..1200
FT /tag= a
FT /product= "GDNF alpha-3"
FT /note= "partial sequence of glial cell-derived
FT neurotrophic factor alpha-3 receptor"

XX
PN EP846764-A2.

XX
PD 10-JUN-1998.

XX
PF 20-NOV-1997; 97EP-00309375.

XX
PR 27-NOV-1996; 96GB-00024677.
PR 09-MAY-1997; 97GB-00009463.

XX
PA (SMIK) SMITHKLINE BEECHAM PLC.

XX
PI Lawrence GMP;

XX
XX WPI; 1998-299980/27.
XX P-PSDB; AAW65116.

XX
PT New factor alpha 3 receptor polypeptide and e.g. DNA and agonists - used
PT to treat neuro degenerative diseases, muscular diseases and nerve and
PT muscle trauma and in diagnostic assays.

XX
PS Claim 9; Fig 1; 22pp; English.

XX
CC This sequence encodes a novel glial cell line-derived neurotrophic factor
CC alpha-3 receptor (GDNF alpha-3). This protein can be used to treat e.g.
CC neurodegenerative diseases (such as Parkinson's Disease, amyotrophic
CC lateral sclerosis (ALS), spinal muscular atrophy (SMA), Huntington's
CC Disease, Alzheimer's Disease, diabetic neuropathy), muscular diseases
CC (including the muscular dystrophies) and nerve and muscle trauma and in
CC diagnostic assays for such conditions

XX
SQ Sequence 1200 BP; 236 A; 419 C; 307 G; 238 T; 0 U; 0 Other;

Query Match 1.9%; Score 22; DB 2; Length 1200;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 808 GCTGCTGCTGCTGCCGCGTCG 829
Db 51 GCTGCTGCTGCTGCCGCGTCG 72

Search completed: September 14, 2004, 23:41:47
Job time : 508.04 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 23:14:24 ; Search time 91.8225 Seconds
(without alignments)
7101.389 Million cell updates/sec

Title: US-09-729-264-1

Perfect score: 1175

Sequence: 1 ctgtctgcccactctgaataa.....gtaatacaactgtagtatag 1175

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1281558

Minimum DB seq length: 16

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PTCUS.COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	1.9	1203	3	US-09-220-528-65
2	22	1.9	1699	4	US-09-187-906-20
3	22	1.9	3942	3	US-09-162-484-19
4	21	1.8	550	4	US-09-669-751-128
5	21	1.8	1428	4	US-09-489-039A-3243
6	20	1.7	394	4	US-09-621-976-16752
7	20	1.7	794	4	US-09-621-976-16751
8	20	1.7	984	4	US-09-252-991A-12897
9	20	1.7	1857	4	US-09-252-991A-12735
10	20	1.7	1910	3	US-09-593-711A-3
11	20	1.7	1914	1	US-07-601-094-1
12	20	1.7	1914	1	US-08-012-735-1
13	20	1.7	2214	3	US-08-864-038A-1
14	20	1.7	3331	3	US-08-864-038A-2
15	20	1.7	3331	3	US-08-864-038A-4
16	20	1.7	6407	2	US-08-616-844-7
17	20	1.7	6407	2	US-08-599-654-7
18	20	1.7	6407	3	US-08-944-868A-7
19	20	1.7	6407	3	US-08-944-423A-7
20	20	1.7	6407	3	US-08-944-496-7
21	19	1.6	402	4	US-09-621-976-88
22	19	1.6	1425	1	US-08-464-148-1
23	19	1.6	1425	1	US-08-385-500-1
24	19	1.6	1425	1	US-08-846-784-1
25	19	1.6	1477	4	US-09-620-312D-1019
26	19	1.6	1768	4	US-09-833-381-523
27	19	1.6	1899	4	US-09-919-060-15

C	28	19	1.6	1899	4	US-09-919-060-16	Sequence 16, Appl
	29	19	1.6	1899	1	US-08-447-500-5	Sequence 5, Appl
	30	19	1.6	1899	1	US-08-454-097-5	Sequence 5, Appl
	31	19	1.6	1899	1	US-08-453-866-5	Sequence 5, Appl
	32	19	1.6	1899	3	US-08-185-359-5	Sequence 5, Appl
	33	19	1.6	1995	4	US-09-620-312D-908	Sequence 908, App
	34	19	1.6	2023	4	US-09-491-522-6	Sequence 6, Appl
C	35	19	1.6	2115	2	US-08-474-379C-60	Sequence 60, Appl
C	36	19	1.6	2115	3	US-09-146-249A-60	Sequence 60, Appl
C	37	19	1.6	2115	3	US-08-206-188B-60	Sequence 60, Appl
	38	19	1.6	2439	4	US-09-489-039A-4904	Sequence 4904, Ap
	39	19	1.6	2450	4	US-09-491-522-2	Sequence 2, Appl
	40	19	1.6	2617	4	US-09-786-240-21	Sequence 21, Appl
	41	19	1.6	2693	4	US-09-919-060-12	Sequence 12, Appl
C	42	19	1.6	2693	4	US-09-919-060-14	Sequence 14, Appl
C	43	19	1.6	3311	4	US-09-367-891A-5	Sequence 5, Appl
C	44	19	1.6	5092	3	US-09-412-545-1	Sequence 1, Appl
	45	19	1.6	5621	4	US-09-566-921-106	Sequence 106, App

ALIGNMENTS

RESULT 1
US-09-220-528-65
; Sequence 65, Application US/09220528A
; Patent No. 6284540
; GENERAL INFORMATION:
; APPLICANT: Milbrandt, Jeffrey D.
; APPLICANT: Baloh, Robert H.
; TITLE OF INVENTION: Artemin, A No. 6284540e1 Neurotrophic Factor
; FILE REFERENCE: 6029-7998
; CURRENT APPLICATION NUMBER: US/09/220,528A
; CURRENT FILING DATE: 1998-12-24
; EARLIER APPLICATION NUMBER: 09/218,698
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 60/108,148
; EARLIER FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/163,283
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 65
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-528-65

Query Match 1.9%; Score 22; DB 3; Length 1203;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 808 GCTGCTGCTGCTGCTGCTGCTGCTG 829
Db 51 GCTGCTGCTGCTGCTGCTGCTGCTG 72

RESULT 2
US-09-187-906-20
; Sequence 20, Application US/09187906
; Patent No. 6677135
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: Ret Ligand (RetL) for Stimulating Neural
; TITLE OF INVENTION: and Renal Growth
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02142

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/187,906
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/07726
FILING DATE: 07-MAY-97
APPLICATION NUMBER: US 60/017,427
FILING DATE: 08-MAY-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/019,300
FILING DATE: 07-JUN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/021,859
FILING DATE: 16-JUL-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/043,533
FILING DATE: 10-APR-97
ATTORNEY/AGENT INFORMATION:
NAME: Kaplan, Warren A.
REGISTRATION NUMBER: 34,199
REFERENCE/DOCKET NUMBER: A008 PCT CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-679-2400
TELEFAX: 617-679-2838
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 1699 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 175..1374
US-09-187-906-20

Query Match 1.9%; Score 22; DB 4; Length 3942;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 808 GCTGCTGCTGCTGCGCGCTCG 829
DB 72 GCTGCTGCTGCTGCGCGCTCG 93

RESULT 4
US-09-669-751-128/c
; Sequence 128, Application US/09669751
; Patent No. 6551575
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Ralph J.
; TITLE OF INVENTION: Methods for Identifying Compounds for
; TITLE OF INVENTION: Motion Sickness, Vertigo and Other Disorders Related to
; TITLE OF INVENTION: Balance and the Perception of Gravity
; FILE REFERENCE: P-NI 3864
; CURRENT APPLICATION NUMBER: US/09/669,751
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 60/168,579
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Drosophila
US-09-669-751-128

Query Match 1.8%; Score 21; DB 4; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 844 ACTGCTGCTGCTGCTGTT 864
DB 317 ACTGCTGCTGCTGCTGTT 297

RESULT 5
US-09-489-039A-3243/c
; Sequence 3243, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709-2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 3243
; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3243

Query Match 1.8%; Score 21; DB 4; Length 1428;
Best Local Similarity 100.0%; Pred. No. 0.58;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GCTTCAACTGCACCGCTCTCCC 141
DB 1147 GCTTCAACTGCACCGCTCTCCC 1127

RESULT 6
US-09-621-976-16752
; Sequence 16752, Application US/09621976
; Patent No. 6639063

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/187,906
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/07726
FILING DATE: 07-MAY-97
APPLICATION NUMBER: US 60/017,427
FILING DATE: 08-MAY-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/019,300
FILING DATE: 07-JUN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/021,859
FILING DATE: 16-JUL-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/043,533
FILING DATE: 10-APR-97
ATTORNEY/AGENT INFORMATION:
NAME: Kaplan, Warren A.
REGISTRATION NUMBER: 34,199
REFERENCE/DOCKET NUMBER: A008 PCT CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-679-2400
TELEFAX: 617-679-2838
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 1699 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 175..1374
US-09-187-906-20

Query Match 1.9%; Score 22; DB 4; Length 1699;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 808 GCTGCTGCTGCTGCGCGCTCG 829
DB 225 GCTGCTGCTGCTGCGCGCTCG 246

RESULT 3
US-09-162-484-19
; Sequence 19, Application US/09162484
; Patent No. 6248724
; GENERAL INFORMATION:
; APPLICANT: Phillips, M. Ian
; APPLICANT: Mohuczy, Dagmara
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS TARGETED TO
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME MRNA AND METHODS OF USE
; FILE REFERENCE: UFLA:087/UFLA087P
; CURRENT APPLICATION NUMBER: US/09/162,484
; CURRENT FILING DATE: 1998-09-25
; EARLIER APPLICATION NUMBER: 60/059,661
; PRIOR FILING DATE: 1997-09-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 3942
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-162-484-19

```
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12897

Query Match      1.7%; Score 20; DB 4; Length 984;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      834 TGTGGCTGCAACTGCTGCTG 853
Db      395 TGTGGCTGCAACTGCTGCTG 414

RESULT 9
US-09-252-991A-12735
; Sequence 12735, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12735
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12735

Query Match      1.7%; Score 20; DB 4; Length 1857;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      834 TGTGGCTGCAACTGCTGCTG 853
Db      1746 TGTGGCTGCAACTGCTGCTG 1765

RESULT 10
US-09-593-711A-3/c
; Sequence 3, Application US/09593711A
; Patent No. 6271030
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Madeline M. Butler
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF C/EBP BETA EXPRESSION
; FILE REFERENCE: RTS-0118
; CURRENT APPLICATION NUMBER: US/09/593,711A
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 244
; SEQ ID NO 3
; LENGTH: 1910
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
; NAME/KEY: unsure
; LOCATION: 1415
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1421
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1422
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1423

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12897

Query Match      1.7%; Score 20; DB 4; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      817 GCTGCCGCGCTGCTGTTGT 836
Db      34 GCTGCCGCGCTGCTGTTGT 53

RESULT 7
US-09-621-976-16751
; Sequence 16751, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16751
; LENGTH: 794
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16751

Query Match      1.7%; Score 20; DB 4; Length 794;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      817 GCTGCCGCGCTGCTGTTGT 836
Db      34 GCTGCCGCGCTGCTGTTGT 53

RESULT 8
US-09-252-991A-12897
; Sequence 12897, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12897
; LENGTH: 984
; TYPE: DNA
```

Wed Sep 15 10:41:06 2004

us-09-729-264-1.1.01g.rni

OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 1424
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 1458
OTHER INFORMATION: unknown
NAME/KEY: CDS
LOCATION: (299)...(1336)
US-09-593-711A-3

Query Match 1.7%; Score 20; DB 3; Length 1910;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 807 CGCTGCTGCTGCTGCGCGCG 826
DB 205 CGCTGCTGCTGCTGCGCGCG 186

RESULT 11

US-07-601-094-1/c
Sequence 1, Application US/07601094
Patent No. 5215892
GENERAL INFORMATION:
APPLICANT: Kishimoto, Tadimitsu
APPLICANT: Hirano, Toshio
APPLICANT: Akira, Shizuo
APPLICANT: Isshiki, Hiroshi
APPLICANT: Tanabe, Osamu
APPLICANT: Kinoshita, Shigemi
APPLICANT: Shimamoto, Takuya
TITLE OF INVENTION: C/EBP2 Gene and Recombinant
TITLE OF INVENTION: C/EBP2
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak &
ADDRESSEE: Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/601,094
FILING DATE: 19901022
CLASSIFICATION: 435
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1914 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 281..1316
OTHER INFORMATION:

Query Match 1.7%; Score 20; DB 1; Length 1914;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 807 CGCTGCTGCTGCTGCGCGCG 826
DB 204 CGCTGCTGCTGCTGCGCGCG 185

RESULT 12

US-08-012-735-1/c
Sequence 1, Application US/08012735
Patent No. 5360894
GENERAL INFORMATION:
APPLICANT: Kishimoto, Tadimitsu
APPLICANT: Hirano, Toshio
APPLICANT: Akira, Shizuo
APPLICANT: Isshiki, Hiroshi
APPLICANT: Tanabe, Osamu
APPLICANT: Kinoshita, Shigemi
APPLICANT: Shimamoto, Takuya
TITLE OF INVENTION: C/EBP2 Gene and Recombinant
TITLE OF INVENTION: C/EBP2
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak &
ADDRESSEE: Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/012,735
FILING DATE: 19930203
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/601,094
FILING DATE: 22 OCT 1990
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1914 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 281..1316
OTHER INFORMATION:

Query Match 1.7%; Score 20; DB 1; Length 1914;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 807 CGCTGCTGCTGCTGCGCGCG 826
DB 204 CGCTGCTGCTGCTGCGCGCG 185

RESULT 13

US-08-864-038A-1
Sequence 1, Application US/08864038A
Patent No. 6001592
GENERAL INFORMATION:
APPLICANT: Kunio NAKASHIMA et al.
TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE cDNA, VECTOR

;; TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
;; TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
;; TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
;; TITLE OF INVENTION: TO SAID POLYPEPTIDE
;; NUMBER OF SEQUENCES: 4
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: 812-5 Hirano
;; STREET: Isshinden
;; CITY: Tsu-city
;; STATE: Mie-prefecture
;; COUNTRY: JAPAN
;; ZIP: 514-01
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: Microsoft Windows 95
;; SOFTWARE: Word Perfect 6.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/864,038A
;; FILING DATE: May 28, 1997
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: JP 8-184459
;; FILING DATE: 15-July-1996
;; NAME: C. Bruce Hamburg
;; REGISTRATION NUMBER: 22,389
;; REFERENCE/DOCKET NUMBER: F-5610
;; TELEPHONE: (212)986-2340
;; TELEFAX: (212)953-7733
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2214
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; ORGANISM: Pinctada fucata
;; CELL TYPE: mantle epithelial cell
US-08-864-038A-1

Query Match 1.7%; Score 20; DB 3; Length 2214;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 807 CGCTGCTGCTGCTGCCGCG 826
|||||
Db 855 CGCTGCTGCTGCTGCCGCG 874
|||||

RESULT 14
US-08-864-038A-2
; Sequence 2, Application US/08864038A
; Patent No. 6001592
; GENERAL INFORMATION:
; APPLICANT: Kunio NAKASHIMA et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
; TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
; TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
; TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 812-5 Hirano
; STREET: Isshinden
; CITY: Tsu-city
; STATE: Mie-prefecture
; COUNTRY: JAPAN
; ZIP: 514-01
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Microsoft Windows 95

;; SOFTWARE: Word Perfect 6.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/864,038A
;; FILING DATE: May 28, 1997
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: JP 8-184459
;; FILING DATE: 15-July-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: C. Bruce Hamburg
;; REGISTRATION NUMBER: 22,389
;; REFERENCE/DOCKET NUMBER: F-5610
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212)986-2340
;; TELEFAX: (212)953-7733
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3331
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA to mRNA
;; ORIGINAL SOURCE:
;; ORGANISM: Pinctada fucata
;; CELL TYPE: mantle epithelial cell
;; FEATURE: mRNA
;; LOCATION: from 1 to 3331
;; IDENTIFICATION METHOD: E (by experiment)
US-08-864-038A-2

Query Match 1.7%; Score 20; DB 3; Length 3331;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 807 CGCTGCTGCTGCTGCCGCG 826
|||||
Db 904 CGCTGCTGCTGCTGCCGCG 923
|||||

RESULT 15
US-08-864-038A-4
; Sequence 4, Application US/08864038A
; Patent No. 6001592
; GENERAL INFORMATION:
; APPLICANT: Kunio NAKASHIMA et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
; TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
; TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
; TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 812-5 Hirano
; STREET: Isshinden
; CITY: Tsu-city
; STATE: Mie-prefecture
; COUNTRY: JAPAN
; ZIP: 514-01
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Microsoft Windows 95
; SOFTWARE: Word Perfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,038A
; FILING DATE: May 28, 1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 8-184459
; FILING DATE: 15-July-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-5610
; TELECOMMUNICATION INFORMATION:

us-09-729-264-1.1.0lig.rni

Wed Sep 15 10:41:06 2004

```

; TELEPHONE: (212)986-2340
; TELEFAX: (212)953-7733
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3331
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Pinctada fucata
; CELL TYPE: mantle epithelial cell
; FEATURE:
; NAME/KEY: CDS
; LOCATION: from 50 to 2263
; IDENTIFICATION METHOD: P (by similarity to some other pattern)
US-08-864-038A-4

Query Match      1.7%; Score 20; DB 3; Length 3331;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      807 CGCTGCTGCTGCTGCTGCCGCCG 826
      |||||||
Db      904 CGCTGCTGCTGCTGCTGCCGCCG 923

Search completed: September 15, 2004, 06:48:26
Job time : 92.8225 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 23:16:34 ; Search time 624.393 Seconds
(without alignments)
9468.459 Million cell updates/sec

Title: US-09-729-264-1

Perfect score: 1175

Sequence: 1 cgtgtgccatcgaataa.....gtaatacaactgtagtatag 1175

Scoring table: OLIGO.NUC

Gapop 60.0 , Gapext 60.0

Searched: 3304383 seqs, 2515761380 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6509420

Minimum DB seq length: 16

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1076	91.6	2051	16	US-10-104-047-1104
2	217	18.5	474	10	US-09-918-995-3342
3	144	12.3	401	9	US-09-864-761-16305
4	25	2.1	357	9	US-09-864-761-16653
5	23	2.0	1284	17	US-10-437-963-81852
6	22	1.9	767	9	US-09-770-445-930
7	22	1.9	1203	9	US-09-220-920-65
8	22	1.9	1829	9	US-09-828-366-15
9	22	1.9	4142	12	US-10-152-319A-1498
10	21	1.8	462	17	US-10-437-963-41506
11	21	1.8	493	10	US-09-918-995-11051
12	21	1.8	550	15	US-10-255-536-128
13	21	1.8	602	17	US-10-767-701-28842
14	21	1.8	1558	10	US-09-930-213-20

C 15	21	1.8	3362	13	US-10-087-192-185
C 16	21	1.8	94317	13	US-10-087-192-184
C 17	20	1.7	330	17	US-10-437-963-55684
C 18	20	1.7	358	13	US-09-823-245A-58
C 19	20	1.7	398	9	US-09-983-965-4945
C 20	20	1.7	412	10	US-09-918-995-5855
C 21	20	1.7	507	17	US-10-430-201-970
C 22	20	1.7	507	17	US-10-430-201-971
C 23	20	1.7	596	17	US-10-767-701-5378
C 24	20	1.7	848	13	US-10-425-114-20418
C 25	20	1.7	863	13	US-10-027-632-150405
C 26	20	1.7	863	16	US-10-027-632-150405
C 27	20	1.7	1192	16	US-10-120-988-90
C 28	20	1.7	1255	13	US-10-425-114-19266
C 29	20	1.7	1290	15	US-10-146-733-16
C 30	20	1.7	1383	17	US-10-437-963-37464
C 31	20	1.7	1526	13	US-10-425-114-31002
C 32	20	1.7	1576	13	US-09-852-386-87
C 33	20	1.7	1814	15	US-10-060-036-182
C 34	20	1.7	1910	9	US-09-789-836-1
C 35	20	1.7	1910	10	US-09-789-831-1
C 36	20	1.7	1910	17	US-10-415-325-18
C 37	20	1.7	1943	15	US-10-146-733-14
C 38	20	1.7	2080	13	US-10-257-022-30
C 39	20	1.7	2504	14	US-10-108-605-246
C 40	20	1.7	2616	17	US-10-437-963-75800
C 41	20	1.7	2874	9	US-09-801-260-3
C 42	20	1.7	2874	15	US-10-145-586-42
C 43	20	1.7	3112	16	US-10-120-988-261
C 44	20	1.7	3226	9	US-09-801-260-1
C 45	20	1.7	3226	15	US-10-145-586-40

ALIGNMENTS

RESULT 1
US-10-104-047-1104
; Sequence 1104, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1104
; LENGTH: 2051
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1104

Query Match	91.6%	Score 1076;	DB 16;	Length 2051;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 1126;	Conservative	0;	Mismatches 1;	Indels 0;
Gaps	0;			
QY	49	ACGGTTCTGGTCTGTAATGAAGTCATAGAGCCGCCAGATGCAAGTCTCTGAAGG	108	
DB	201	ACGGTTCTGGTCTGTAATGAAGTCATAGAGCCGCCAGATGCAAGTCTCTGAAGG	260	
QY	109	GCTCCAGGCTCGCTTCAACTGCACCGTCTCCAGGGCTGGAAGTCATCATGTGGGCTC	168	
DB	261	GCTCCAGGCTCGCTTCAACTGCACCGTCTCCAGGGCTGGAAGTCATCATGTGGGCTC	320	
QY	169	TAGTGACATGTGGTGTAAAGCGTCAGGCCCATCATCAATGACCGGT	228	
DB	321	TAGTGACATGTGGTGTAAAGCGTCAGGCCCATCATCAATGACCGGT	380	
QY	229	TCACCTCTCAGAGGTACGACACAGGGGGGAATTCTACCTCGGAGATGATCATCCCAATG	288	

Db	381	TCACCTCTCAGAGGTACGACGGCGGAACTTCACCTCGAGATGATCATCCACAATG	440
Qy	289	TGGAGCCAGTGATTCGGGAAACATCAGATGCAGCCTCCAGAACAGTGCCTGCATGAT	348
Db	441	TGGAGCCAGTGATTCGGGAAACATCAGATGCAGCCTCCAGAACAGTGCCTGCATGAT	500
Qy	349	CTGCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTCAATCCCAAGTGTAACTTTGAT	408
Db	501	CTGCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTCAATCCCAAGTGTAACTTTGAT	560
Qy	409	TCGCTGAGAATGAACCTTGTGAAGTTACTTGTCTACCTCACAATGGAACCGGCTCCGG	468
Db	561	TCGCTGAGAATGAACCTTGTGAAGTTACTTGTCTACCTCACAATGGAACCGGCTCCGG	620
Qy	469	ATATTTCTCGGAGCTCGGTCTCTGTGTCAGCCATTCAAGCTATTATTTTGTTCGGAGC	528
Db	621	ATATTTCTCGGAGCTCGGTCTCTGTGTCAGCCATTCAAGCTATTATTTTGTTCGGAGC	680
Qy	529	CCAGCGACCTTCAAAGTGCAGTGAGCAATCTGGCTCGACCCACAGAGCAATGGGACTT	588
Db	681	CCAGCGACCTTCAAAGTGCAGTGAGCAATCTGGCTCGACCCACAGAGCAATGGGACTT	740
Qy	589	TGACTTGGCTCGCTACCTGGAAGAGCTGGAAGCCCGCAAGTCTGCAACTGTAAATCTCA	648
Db	741	TGACTTGGCTCGCTACCTGGAAGAGCTGGAAGCCCGCAAGTCTGCAACTGTAAATCTCA	800
Qy	649	CTGTGATTCGGTGTCCCCCAAGCACTGGAGGTGGTATTAATATTTCCAGSTGTATTATCAA	708
Db	801	CTGTGATTCGGTGTCCCCCAAGCACTGGAGGTGGTATTAATATTTCCAGSTGTATTATCAA	860
Qy	709	GTTTACCAGATTAGTGTGTTTTCAATTCCTACTTCGGGCAAGTTGAGTTGGACTAGCAG	768
Db	861	GTTTACCAGATTAGTGTGTTTTCAATTCCTACTTCGGGCAAGTTGAGTTGGACTAGCAG	920
Qy	769	GCACCATGCTTCTGACGGCAGCTGACTCTTACAAATACGCTGCTGCTGCGCGCGCTC	828
Db	921	GCACCATGCTTCTGACGGCAGCTGACTCTTACAAATACGCTGCTGCTGCGCGCGCTC	980
Qy	829	GTTTGTGTGGCTGCAACTGCTGCTGCCGTGTTGTTCTGCTGAGAAGAAAAGAGGAT	888
Db	981	GTTTGTGTGGCTGCAACTGCTGCTGCCGTGTTGTTCTGCTGAGAAGAAAAGAGGAT	1040
Qy	889	TTTCGTTATCAATTTCAAAGAATACTGAAAAGAGAGACACAAACAAAGAACTGAGACAG	948
Db	1041	TTTCGTTATCAATTTCAAAGAATACTGAAAAGAGAGACACAAACAAAGAACTGAGACAG	1100
Qy	949	AAAGTGGAAATGAAATCCCGGCTACAAATTCAGATGAAACAAAGACACACAGACCGCTT	1008
Db	1101	AAAGTGGAAATGAAATCCCGGCTACAAATTCAGATGAAACAAAGACACACAGACCGCTT	1160
Qy	1009	CTCTCCCTCCAAATCCTGTGTAATCCAGTGATCTCTGAACAAAGAAACAGTAGCTGTGCC	1068
Db	1161	CTCTCCCTCCAAATCCTGTGTAATCCAGTGATCTCTGAACAAAGAAACAGTAGCTGTGCC	1220
Qy	1069	CTCCTCACAGCGGGCTGATCAAGCTCACCCAGCGCCAGCAAGTCATCCACAGGCTTCTT	1128
Db	1221	CTCCTCACAGCGGGCTGATCAAGCTCACCCAGCGCCAGCAAGTCATCCACAGGCTTCTT	1280
Qy	1129	TTAATCTGCCAGTCTCTGAGAAGGTGAGTAATATCAACTGTGTAGTATAG	1175
Db	1281	TTAATCTGCCAGTCTCTGAGAAGGTGAGTAATATCAACTGTGTAGTATAG	1327

RESULT 2

US-09-918-995-3342
Sequence 3342, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID
; TITLE OF INVENTION: FROM VARIOUS GDN
; FILE REFERENCE: 20411-756

;	CURRENT APPLICATION NUMBER:	US/09/918,995	
;	CURRENT FILING DATE:	2001-07-30	
;	PRIOR APPLICATION NUMBER:	US/09/235,076	
;	PRIOR FILING DATE:	1999-01-20	
;	NUMBER OF SEQ ID NOS:	38054	
;	SOFTWARE:	FastSeq for Windows Version 3.0	
;	SEQ ID NO 3342		
;	LENGTH:	474	
;	TYPE:	DNA	
;	ORGANISM:	Homo sapiens	
;	FEATURE:		
;	NAME/KEY:	misc feature	
;	LOCATION:	(1)..(474)	
;	OTHER INFORMATION:	n = A,T,C or G	
;	US-09-918-995-3342		

Query Match	18.5%;	Score 217;	DB 10;	Length 474;
Best Local Similarity	99.6%;	Pred.No.2.3e-102;		
Matches 267;	Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;

Qy	908	GAATCTGAAAAGAGAGAACAAACAAGAAACTGTGACAGAGAAAGTGGAAATGAAAACTC	967
Db	63	GAATCTGAAAAGAGAGAAAGCAACAAGAAACTGTGACAGAGAAAGTGGAAATGAAAACTC	122
Qy	968	CGGCTACAAATTCAGATGAAACAAAAGACACAGACACCGCTTCTCCTCCCAAAATCCTG	1027
Db	123	CGGCTACAAATTCAGATGAAACAAAAGACACAGAAACCGCTTCTCCTCCCAAAATCCTG	182
Qy	1028	TGAATCAGTAGTATCTGAACAAAGAAACAGTAGCTGTGGCCCTCCTCACACAGCGGCTGA	1087
Db	183	TGAATCAGTAGTATCTGTAAACAAAGAAACAGTAGTGTGGCCCTCCTCACACAGCGGCTGA	242
Qy	1088	TCAAGTCCACCCAGGCGCAGCAAGTCAATCCACAGGCTTCTTTTAAATCTGGCAGTCTCTGA	1147
Db	243	TCAAGTCCACCCAGGCGCAGCAAGTCAATCCACAGGCTTCTTTTAAATCTGGCAGTCTCTGA	302
Qy	1148	GAAGTTCAGTAATCAACTGTAGTATAG	1175
Db	303	GAAGTTCAGTAATCAACTGTAGTATAG	330

RESULT 3	
US-09-864-761-16305	
Sequence 16305, Application US/09864761	
Patent No. US20020048763A1	
GENERAL INFORMATION:	
APPLICANT: Penn, Sharon G.	
APPLICANT: Rank, David R.	
APPLICANT: Hanzel, David K.	
APPLICANT: Chen, Wensheng	
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES US	
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY	
FILE REFERENCE: Aecomica-X-1	
CURRENT APPLICATION NUMBER: US/09/864,761	
CURRENT FILING DATE: 2001-05-23	
PRIOR APPLICATION NUMBER: US 60/180,312	
PRIOR FILING DATE: 2000-02-04	
PRIOR APPLICATION NUMBER: US 60/207,456	
PRIOR FILING DATE: 2000-05-26	
PRIOR APPLICATION NUMBER: US 09/632,366	
PRIOR FILING DATE: 2000-08-03	
PRIOR APPLICATION NUMBER: GB 24263.6	
PRIOR FILING DATE: 2000-10-04	
PRIOR APPLICATION NUMBER: US 60/236,359	
PRIOR FILING DATE: 2000-09-27	
PRIOR APPLICATION NUMBER: PCT/US01/00666	
PRIOR FILING DATE: 2001-01-30	
PRIOR APPLICATION NUMBER: PCT/US01/00667	
PRIOR FILING DATE: 2001-01-30	
PRIOR APPLICATION NUMBER: PCT/US01/00664	
PRIOR FILING DATE: 2001-01-30	
PRIOR APPLICATION NUMBER: PCT/US01/00669	
PRIOR FILING DATE: 2001-01-30	

```
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16305
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF121782.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
US-09-864-761-16305

Query Match      12.3%; Score 144; DB 9; Length 401;
Best Local Similarity 100.0%; Pred. No. 3e-64;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 AGTCCTGAAGGGCTCCAGGCTCTCAACTGCACCGTCTCCAGGCTGGAGCTCAT 157
Db 258 AGTCCTGAAGGGCTCCAGGCTCTCAACTGCACCGTCTCCAGGCTGGAGCTCAT 317

QY 158 CATGTGGGCTCTCAGTGACATGTGTGCTTAAGCGTCAGGCCCATGAGCCCATCATCAC 217
Db 318 CATGTGGGCTCTCAGTGACATGTGTGCTTAAGCGTCAGGCCCATGAGCCCATCATCAC 377

QY 218 CAATGACCGCTTCACCTCTCAGAG 241
Db 378 CAATGACCGCTTCACCTCTCAGAG 401

RESULT 4
US-09-864-761-16653
; Sequence 16653, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16653
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF064857.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.93
US-09-864-761-16653

Query Match      2.1%; Score 25; DB 9; Length 357;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 884 AGGATTTCGTATTCAATTTCAAAAG 908
Db 42 AGGATTTCGTATTCAATTTCAAAAG 66

RESULT 5
US-10-437-963-81852/c
; Sequence 81852, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 81852
; LENGTH: 1284
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MFT4530_81337C.1
US-10-437-963-81852

Query Match      2.0%; Score 23; DB 17; Length 1284;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Wed Sep 15 10:41:06 2004

```

; SEQ ID NO 65
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-920-65

Query Match      1.9%; Score 22; DB 9; Length 1203;
Best Local Similarity 100.0%; Pred. No. 1.5; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY      808 GCTGCTGCTGCTGCGCGCGTCG 829
Db      51 GCTGCTGCTGCTGCGCGCGTCG 72

RESULT 8
US-09-828-366-15
; Sequence 15, Application US/09828366
; Patent No. US20020010137A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Klein, Robert D.
; APPLICANT: Napier, Mary
; APPLICANT: Wood, William I.
; APPLICANT: Yuan, Jean
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC
; TITLE OF INVENTION: CELL GROWTH
; FILE REFERENCE: P1694R1C1
; CURRENT APPLICATION NUMBER: US/09/828,366
; CURRENT FILING DATE: 2001-04-05
; Prior filing data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 29
; SEQ ID NO 15
; LENGTH: 1829
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-828-366-15

Query Match      1.9%; Score 22; DB 9; Length 1829;
Best Local Similarity 100.0%; Pred. No. 1.5; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY      808 GCTGCTGCTGCTGCGCGCGTCG 829
Db      88 GCTGCTGCTGCTGCGCGCGTCG 109

RESULT 9
US-10-152-319A-1498
; Sequence 1498, Application US/10152319A
; Publication No. US20040072160A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,910
; PRIOR FILING DATE: 2001-07-10
```

```

; SEQ ID NO 930
; LENGTH: 767
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-445-930

Query Match      1.9%; Score 22; DB 9; Length 767;
Best Local Similarity 100.0%; Pred. No. 1.5; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY      821 CCGCGCTGCTGTTGTGCTGTC 842
Db      672 CCGCGCTGCTGTTGTGCTGTC 693

RESULT 7
US-09-220-920-65
; Sequence 65, Application US/09220920
; Patent No. US2002002269A1
; GENERAL INFORMATION:
; APPLICANT: Milbrandt, Jeffrey D.
; APPLICANT: Balch, Robert H.
; TITLE OF INVENTION: Artemin, A No. US20020002269A1el Neurotrophic Factor
; FILE REFERENCE: 6029-7996
; CURRENT APPLICATION NUMBER: US/09/220,920
; CURRENT FILING DATE: 1998-12-24
; EARLIER APPLICATION NUMBER: 09/163,283
; EARLIER FILING DATE: 1998-09-29
; EARLIER APPLICATION NUMBER: 60/108,148
; EARLIER FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/218,698
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 2.0
```

```
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1498
; LENGTH: 4142
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_012544
US-10-152-319A-1498

Query Match
Best Local Similarity 1.9%; Score 22; DB 12; Length 4142;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 808 GCTGCTGCTGCGCGCGTCG 829
Db 95 GCTGCTGCTGCGCGCGTCG 116

RESULT 10
US-10-437-963-41506/c
; Sequence 41506, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 41506
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_44844C.1
US-10-437-963-41506

Query Match
Best Local Similarity 1.8%; Score 21; DB 17; Length 462;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 809 CTGCTGCTGCTGCGCGTCG 829
Db 133 CTGCTGCTGCTGCGCGTCG 113

RESULT 11
US-09-918-995-11051/c
; Sequence 11051, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
```

```
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11051
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(493)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-11051

Query Match
Best Local Similarity 1.8%; Score 21; DB 10; Length 493;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 811 GCTGCTGCTGCGCGCGTCGTT 831
Db 443 GCTGCTGCTGCGCGCGTCGTT 423

RESULT 12
US-10-255-536-128/c
; Sequence 128, Application US/10255536
; Publication No. US20030087807A1
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Ralph J.
; TITLE OF INVENTION: Methods for Identifying Compounds for
; TITLE OF INVENTION: Motion Sickness, Vertigo and Other Disorders Related to
; TITLE OF INVENTION: Balance and the Perception of Gravity
; FILE REFERENCE: P-NI 3864
; CURRENT APPLICATION NUMBER: US/10/255,536
; CURRENT FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: US/09/669,751
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 60/168,579
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Drosophila
; FEATURE:
; OTHER INFORMATION:
US-10-255-536-128

Query Match
Best Local Similarity 1.8%; Score 21; DB 15; Length 550;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 844 ACTGCTGCTGCCCGTTGTTGTT 864
Db 317 ACTGCTGCTGCCCGTTGTTGTT 297

RESULT 13
US-10-767-701-28842/c
; Sequence 28842, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
```

;; CURRENT APPLICATION NUMBER: US/10/767,701
;; CURRENT FILING DATE: 2004-01-29
;; NUMBER OF SEQ ID NOS: 63128
;; SEQ ID NO 28842
;; LENGTH: 602
;; TYPE: DNA
;; ORGANISM: Sorghum bicolor
;; FEATURE:
;; OTHER INFORMATION: Clone ID: 8088472
US-10-767-701-28842

Query Match 1.8%; Score 21; DB 17; Length 602;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 810 TGCTGCTGCGCGCGTGT 830
Db 206 TGCTGCTGCGCGCGTGT 186

RESULT 14

US-09-930-213-20/c
;; Sequence 20, Application US/09930213
;; Publication No. US20030170625A1
;; GENERAL INFORMATION:
;; APPLICANT: ROSENTHAL, ANDRE
;; APPLICANT: HINZMANN, BERND
;; APPLICANT: SCHAFER, REINHARD
;; APPLICANT: ZUBER, JOHANNES
;; APPLICANT: TCHE-NITSE, OLEG
;; APPLICANT: GRIPS, MARTIN
;; APPLICANT: HELLMER, MARTIN
;; APPLICANT: SCHMITZ, ANNE-CHANTAL
;; APPLICANT: SERS, CHRISTINE
;; TITLE OF INVENTION: DETECTION OF DIFFERENTIAL GENE EXPRESSIONS
;; FILE REFERENCE: ALBRE-14
;; CURRENT APPLICATION NUMBER: US/09/930,213
;; CURRENT FILING DATE: 2001-01-31
;; PRIOR APPLICATION NUMBER: DE 10004102.7
;; PRIOR FILING DATE: 2000-01-31
;; NUMBER OF SEQ ID NOS: 885
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 20
;; LENGTH: 1558
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (16)
;; OTHER INFORMATION: a, t, c, g, other or unknown
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (427)
;; OTHER INFORMATION: a, t, c, g, other or unknown
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (492)
;; OTHER INFORMATION: a, t, c, g, other or unknown
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (551)
;; OTHER INFORMATION: a, t, c, g, other or unknown
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (728)
;; OTHER INFORMATION: a, t, c, g, other or unknown
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (744)
;; OTHER INFORMATION: a, t, c, g, other or unknown
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (979)

;; OTHER INFORMATION: a, t, c, g, other or unknown
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (1146)
;; OTHER INFORMATION: a, t, c, g, other or unknown
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (1243)
;; OTHER INFORMATION: a, t, c, g, other or unknown
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (1343)
;; OTHER INFORMATION: a, t, c, g, other or unknown
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (1346)
;; OTHER INFORMATION: a, t, c, g, other or unknown
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (1352)
;; OTHER INFORMATION: a, t, c, g, other or unknown
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (1423)
;; OTHER INFORMATION: a, t, c, g, other or unknown
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (1535)
;; OTHER INFORMATION: a, t, c, g, other or unknown
US-09-930-213-20

Query Match 1.8%; Score 21; DB 10; Length 1558;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 808 GCTGCTGCTGCTGCGCGTC 828
Db 223 GCTGCTGCTGCTGCGCGTC 203

RESULT 15

US-10-087-192-185/c
;; Sequence 185, Application US/10087192
;; Publication No. US20020182586A1
;; GENERAL INFORMATION:
;; APPLICANT: Morris, David W.
;; APPLICANT: Engelhard, Eric K.
;; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;; TITLE OF INVENTION: CANCER
;; FILE REFERENCE: 529452000122
;; CURRENT APPLICATION NUMBER: US/10/087,192
;; CURRENT FILING DATE: 2002-03-01
;; PRIOR APPLICATION NUMBER: US 09/747,377
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: US 09/798,586
;; PRIOR FILING DATE: 2001-03-02
;; NUMBER OF SEQ ID NOS: 2059
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 185
;; LENGTH: 3362
;; TYPE: DNA
;; ORGANISM: Mus musculus
US-10-087-192-185

Query Match 1.8%; Score 21; DB 13; Length 3362;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 808 GCTGCTGCTGCTGCGCGTC 828
Db 947 GCTGCTGCTGCTGCGCGTC 927

Wed Sep 15 10:41:06 2004

us-09-729-264-1.olig.rnpb

Page 7

Search completed: September 15, 2004, 07:20:16
Job time : 626.393 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 20:48:25 ; Search time 4800.11 Seconds
(without alignments)
10546.560 Million cell updates/sec

Title: US-09-729-264-3
Perfect score: 1168
Sequence: 1 agtgatcatgttgcaggag.....gtaatacaactgtagtagatag 1168

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6784142

Minimum DB seq length: 16
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_ste.*

28: em_un.*

29: em_vi.*

30: em_hg_hum.*

31: em_hg_inv.*

32: em_hg_other.*

33: em_hg_mus.*

34: em_hg_pin.*

35: em_hg_rtd.*

36: em_hg_mam.*

37: em_hg_vrt.*

38: em_sy.*

39: em_hggo_hum.*

40: em_hggo_mus.*

41: em_hggo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1168	100.0	1168	6	AX380398	AX380398 Sequence
2	1027	87.9	1175	6	AX380396	AX380396 Sequence
3	976	83.6	1240	6	AX380400	AX380400 Sequence
4	976	83.6	2051	6	AX747579	AX747579 Sequence
5	976	83.6	2051	9	AK092316	AK092316 Homo sapi
6	771	66.0	1139	6	AX380402	AX380402 Sequence
7	319	27.3	142742	9	AF121782	AF121782 Homo sapi
8	319	27.3	340000	9	HS21C080	AL163280 Homo sapi
9	241	20.6	182532	9	CH179K04	AL954228 Pan trogl
10	145	12.4	156288	9	BS000162	BS000162 Pan trogl
11	100	8.6	199665	9	AF064857	AF064857 Homo sapi
12	72	5.2	192219	9	RF43002119	AL954227 Pan trogl
13	40	3.4	170121	9	AF064860	AF064860 Homo sapi
14	25	2.1	6470	3	AB090820	AB090820 Anopheles
15	24	2.1	5178	3	AY119603	AY119603 Drosophila
16	24	2.1	8152	3	DMTOC	Y14157 Drosophila
17	24	2.1	77137	3	AC004422	AC004422 Drosophila
18	24	2.1	83876	2	AC020009	AC020009 Drosophila
19	24	2.1	148102	8	AF003328	AP003328 Oryza sat
20	24	2.1	148762	8	AF002843	AP002843 Oryza sat
21	24	2.1	169931	3	AC008321	AC008321 Drosophila
22	24	2.1	184554	3	AC099022	AC099022 Drosophila
23	24	2.1	268984	3	AE001274	AE001274 Leishmani
24	24	2.1	314957	3	AE003581	AE003581 Drosophila
25	23	2.0	1155	10	AF537215	AF537215 Mus muscu
26	23	2.0	1195	6	AX380404	AX380404 Sequence
27	23	2.0	29865	2	AC020047	AC020047 Drosophila
28	23	2.0	106256	3	AC108135	AC108135 Leishmani
29	23	2.0	124612	2	AC105442	AC105442 Leishmani
30	23	2.0	149964	2	AC120145	AC120145 Mus muscu
31	23	2.0	166000	3	AC104511	AC104511 Drosophila
32	23	2.0	167926	3	AC023685	AC023685 Drosophila
33	23	2.0	181510	2	AC120346	AC120346 Mus muscu
34	23	2.0	186030	8	CNS08C7W	AL731747 Oryza sat
35	23	2.0	196900	2	AC020851	AC020851 Mus muscu
36	23	2.0	332029	3	AE003491	AE003491 Drosophila
37	22	1.9	646	8	BT005060	BT005060 Arabidops
38	22	1.9	757	8	BT004026	BT004026 Arabidops
39	22	1.9	1200	6	A92047	A92047 Sequence 1
40	22	1.9	1200	6	A92049	A92049 Sequence 3
41	22	1.9	1203	6	BD251402	BD251402 Novel neu
42	22	1.9	1203	6	AX658358	AX658358 Sequence
43	22	1.9	1203	9	AF051767	AF051767 Homo sapi
44	22	1.9	1699	6	BD195081	BD195081 Ret ligan
45	22	1.9	1792	6	BD137012	BD137012 GFR alpha

ALIGNMENTS

RESULT 1
AX380398
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AX380398
Sequence 3 from Patent WO0200710.
AX380398
AX380398.1 GI:19575328
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
B7-like molecules and uses thereof
Patent: WO 0200710-A 3 03-JAN-2002;

AX380398
Sequence 3 from Patent WO0200710.
AX380398
AX380398.1 GI:19575328
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
B7-like molecules and uses thereof
Patent: WO 0200710-A 3 03-JAN-2002;

Pred. No. is the number of results predicted by chance to have a

Amgen, Inc. (US)		Location/Qualifiers	
FEATURES		1..1168	
source		/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
CDS		8..1168	
		/note="unnamed protein product"	
		/codon_start=1	
		/protein_id="CAD28217.1"	
		/db_xref="GI:19575329"	
		/db_xref="REMTREMBL:CAD28217"	
		/translation="WVAGAMENRDPGSGSNEVIEGPONARVLKGSQARFNCTVSGQWKLINWLSVVRMEPIITNDRTSORYDOGNFTSEMIHNHVPDSGNIRCSLONSLRHSAYLTQVVGELFIPSVNLVAENECVETCLPSHWTLNLPDIWELGLLVSHSYFVPEPSDLSQSAVSIILALTPQSNGLTVCATWKSARKASATVNLIVIRCPQDTGGINIPGVLSLPSLFTGWKVLGLAGTMLTPTCTLIIRCCCHRRCCGCCNCCRCCKRKRFRIOFKSEKTKETETESGNSGNSYNSDEQKTTDTASLPFKSCSSDPEQRNSSCGPPHQADQRP RPASHPQASFNLASPEKVSNTTVV"	
ORIGIN			
Query Match		100.0%; Score 1168; DB 6; Length 1168;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 1168; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	AGTGATCATGTGGCAGGAGCCATGGAATAATAGACACCCCGGTTCTGGGTCTGGTAA	60
Db	1	AGTGATCATGTGGCAGGAGCCATGGAATAATAGACACCCCGGTTCTGGGTCTGGTAA	60
QY	61	TGAAGTCATAGAGGCCCCCAAAATCAAGAGTCCTGAAGGGCTCCAGGCTCGCTCAA	120
Db	61	TGAAGTCATAGAGGCCCCCAAAATCAAGAGTCCTGAAGGGCTCCAGGCTCGCTCAA	120
QY	121	CTGCACCGTCTCCAGGCTGGAAGTCATCATGTGGGCTCTCAGTGACATGGTGGCT	180
Db	121	CTGCACCGTCTCCAGGCTGGAAGTCATCATGTGGGCTCTCAGTGACATGGTGGCT	180
QY	181	AAGCGTCAGGCCATGAGGCCCATCATCAATGACCGCTTCACTCTCAGAGGTACGA	240
Db	181	AAGCGTCAGGCCATGAGGCCCATCATCAATGACCGCTTCACTCTCAGAGGTACGA	240
QY	241	CCAGGCGGGAACCTTCACTCGGAGATGATCATCAATGTGAGGCCAGTGTGCTGGG	300
Db	241	CCAGGCGGGAACCTTCACTCGGAGATGATCATCAATGTGAGGCCAGTGTGCTGGG	300
QY	301	GAAATCAGATGACGCTCCAGAACAGTGGCTGATGATGATGATGATGATGATGATG	360
Db	301	GAAATCAGATGACGCTCCAGAACAGTGGCTGATGATGATGATGATGATGATGATG	360
QY	361	AGTTATGGAGAGCTGTTCACTCCAGTGTAACTTGTACTCGTGAATGAACCTTG	420
Db	361	AGTTATGGAGAGCTGTTCACTCCAGTGTAACTTGTACTCGTGAATGAACCTTG	420
QY	421	TGAAGTTACTTGTCTACCTCCACTGACCTGGCTCCCGGATTTCTCGGAGCTCGG	480
Db	421	TGAAGTTACTTGTCTACCTCCACTGACCTGGCTCCCGGATTTCTCGGAGCTCGG	480
QY	481	TCTCTGTGTAGCCATTCAAGCTATATTGTTTCCGAGGCCAGCGACCTTCAAGTGC	540
Db	481	TCTCTGTGTAGCCATTCAAGCTATATTGTTTCCGAGGCCAGCGACCTTCAAGTGC	540
QY	541	AGTGAGCATCTGGCTCTGACCCACAGGCAATGGATTTGACTTGGTGGCTACCTG	600
Db	541	AGTGAGCATCTGGCTCTGACCCACAGGCAATGGATTTGACTTGGTGGCTACCTG	600
QY	601	GAAGAGCTTGAAGGCCCGCAAGTCTGCAACTGTAATCTCACTGTGANTCGGTGCCA	660
Db	601	GAAGAGCTTGAAGGCCCGCAAGTCTGCAACTGTAATCTCACTGTGANTCGGTGCCA	660
QY	661	AGACACTGAGGTGGTATTAATTAATCCAGGTGATTAATCAAGTTTACCGATTAGTTT	720
Db	661	AGACACTGAGGTGGTATTAATTAATCCAGGTGATTAATCAAGTTTACCGATTAGTTT	720

QY	721	TTCAATGCTACTTGGGGCAAAGTTGGACTTGGACTAGCAGGCACCATGCTTCTGACGCC	780
Db	721	TTCAATGCTACTTGGGGCAAAGTTGGACTTGGACTAGCAGGCACCATGCTTCTGACGCC	780
QY	781	GACGTGACTCTTACAATACGCTGCTGCTGCGCGCTGCTTGTGCTGCAACTG	840
Db	781	GACGTGACTCTTACAATACGCTGCTGCTGCGCGCTGCTTGTGCTGCAACTG	840
QY	841	CTGCTGCGCTGTTGTTTCTGTAGAGAAAAGAGGATTCGATTTCAATTTCAAAA	900
Db	841	CTGCTGCGCTGTTGTTTCTGTAGAGAAAAGAGGATTCGATTTCAATTTCAAAA	900
QY	901	GAATCTGAAAAGAGAGAGACAAACAAAGAACTGAGACAGAAAGTGAAAATCAAACTC	960
Db	901	GAATCTGAAAAGAGAGAGACAAACAAAGAACTGAGACAGAAAGTGAAAATCAAACTC	960
QY	961	CGGCTACAAATTCAGATGAACAAAAGACCAACAGACCGCTTCTTCCCTCCAAATCTCTG	1020
Db	961	CGGCTACAAATTCAGATGAACAAAAGACCAACAGACCGCTTCTTCCCTCCAAATCTCTG	1020
QY	1021	TGAATCCAGTGATCTGAAACAAAGAAACAGTAGCTGTGGCCCTCTCACACGGGCTGA	1080
Db	1021	TGAATCCAGTGATCTGAAACAAAGAAACAGTAGCTGTGGCCCTCTCACACGGGCTGA	1080
QY	1081	TCAAGCTCCACCCAGCCAGCAAGTCATCCACAGGCTTCTTTAAATCTGCCAGTCTCTGA	1140
Db	1081	TCAAGCTCCACCCAGCCAGCAAGTCATCCACAGGCTTCTTTAAATCTGCCAGTCTCTGA	1140
QY	1141	GAAAGTCAGTAATACAACTGTAGTATAG	1168
Db	1141	GAAAGTCAGTAATACAACTGTAGTATAG	1168

AX380396	1175 bp	DNA	linear	PAT 18-MAR-2002							
Sequence 1 from Patent WO0200710.											
AX380396	GI:19575326										
AX380396.1											
KEYWORDS	Homo sapiens (human)										
SOURCE	Homo sapiens										
ORGANISM	Homo sapiens										
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.											
REFERENCE	1										
AUTHORS	Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.										
TITLE	B7-like molecules and uses thereof										
JOURNAL	Patent: WO 0200710-A 1 03-JAN-2002;										
Amgen, Inc. (US)											
FEATURES	Location/Qualifiers										
1..1175											
/organism="Homo sapiens"											
/mol_type="unassigned DNA"											
/db_xref="taxon:9606"											
27..1175											
/note="unnamed protein product"											
/codon_start=1											
/protein_id="CAD28216.1"											
/db_xref="GI:19575327"											
/translation="WGLVFLHSGSNEVIEGPONATVLRKGSQARFNCTVSGQWKLINWLSDMVLSVRMEPIITNDRTSORYDOGNFTSEMIHNHVPDSGNIRCSLQNSRLHGSALTYQVVGELFIPSVNLVAENECVETCLPSHWTRLPDIWELGLLVSHSYFVPEPSDLSQSAVSIILALTPQSNGLTVCATWKSARKASATVNLIVIRCPQDGTINIPGVLSLPSLFTGWKVLGLAGTMLTPTCTLIIRCCCHRRCCGCCNCCRCCKRKRFRIOFKSEKTKETETESGNSGNSYNSDEQKTTDTASLPFKSCSSDPEQRNSSCGPPHQADQRP RPASHPQASFNLASPEKVSNTTVV"											

Query Match	87.9%;	Score 1027;	DB 6;	Length 1175;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1077; Conservative	0; Mismatches	1; Indels	0; Gaps	0;

ORIGIN			
--------	--	--	--

LOCUS AK092516 2051 bp mRNA linear PRI 15-JUL-2002
 DEFINITION Homo sapiens cDNA FLJ35197 fis, clone PLACE6017788, highly similar
 to IGSF5.
 ACCESSION AK092516
 VERSION AK092516.1 GI:21751130
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1
 Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,
 Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
 Kawai-Hio, Y., Saito, K., Nishikawa, I., Kimura, K., Yamashita, H.,
 Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
 Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
 Oshima, A., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K.
 and Isogai, T.
 NEDO human cDNA sequencing project
 UNPUBLISHED
 2 (bases 1 to 2051)
 Isogai, T. and Yamamoto, J.
 Direct Submission
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.
 FEATURES
 source
 1. 2051
 /location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE6017788"
 /tissue_type="placenta"
 /clone_lib="PLACE6"
 /note="Cloning vector: pME18SFL3"
 ORIGIN
 Query Match 83.6%; Score 976; DB 9; Length 2051;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1076; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 91 AGTCCTGAGGGCTCCAGGCTCGCTTCAACTGACCGCTCCAGGGCTGGAAGCTCAT 150
 DB 250 AGTCCTGAGGGCTCCAGGCTCGCTTCAACTGACCGCTCCAGGGCTGGAAGCTCAT 309
 QY 151 CATGTGGGCTCTCAGTGACATGTGTGTGCTTAAGCGTCAGGCCATGAGGCCCATCATCAC 210
 DB 310 CATGTGGGCTCTCAGTGACATGTGTGTGCTTAAGCGTCAGGCCATGAGGCCCATCATCAC 369
 QY 211 CAATGACCGCTTCACTCTCAGAGGTACGACACGAGGGGGGAAGCTTCACTCGGAGATGAT 270
 DB 370 CAATGACCGCTTCACTCTCAGAGGTACGACACGAGGGGGGAAGCTTCACTCGGAGATGAT 429
 QY 271 CATCCACATGTGGAGCCAGTGATTTGGGGAAACATCAGATGACGCTCCAGAACAGTCG 330
 DB 430 CATCCACATGTGGAGCCAGTGATTTGGGGAAACATCAGATGACGCTCCAGAACAGTCG 489
 QY 331 CTTGATGATCTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 390
 DB 490 CTTGATGATCTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 549
 QY 391 TAATCTTGTAGTGGTGAAGTGAACCTTGTGAAGTTACTTGTACCTTACCTTACCTTACCTTACCT 450
 DB 550 TAATCTTGTAGTGGTGAAGTGAACCTTGTGAAGTTACTTGTACCTTACCTTACCTTACCTTACCT 609

QY 451 CTGGCTCCCGGATATTTCTGGAGCTCGTCTCTCTGGTCAAGCTATTTATTT 510
 DB 610 CGGGCTCCCGGATATTTCTGGAGCTCGTCTCTCTGGTCAAGCTATTTATTT 669
 QY 511 TGTTCGGAGCCAGCGACCTTCAAAGTGCAGTGAGCATCTTGGCTCTGACCCACAGAG 570
 DB 670 TGTTCGGAGCCAGCGACCTTCAAAGTGCAGTGAGCATCTTGGCTCTGACCCACAGAG 729
 QY 571 CAATGGGATTTGACTTGGTGGCTTCACTGGAGAGCCCTGAAGGCCGCAAGTCTGCAAC 630
 DB 730 CAATGGGATTTGACTTGGTGGCTTCACTGGAGAGCCCTGAAGGCCGCAAGTCTGCAAC 789
 QY 631 TGTAAATCTCACTGTGATTCGGTGTCCCAAGACACTGGAGTGTGTTAATTAATTCAGG 690
 DB 790 TGTAAATCTCACTGTGATTCGGTGTCCCAAGACACTGGAGTGTGTTAATTAATTCAGG 849
 QY 691 TGTATTAATCAAGTTTACCGAGTTAGGTTTTCATTGCTCTCTGGGCAAGTTGGACT 750
 DB 850 TGTATTAATCAAGTTTACCGAGTTAGGTTTTCATTGCTCTCTGGGCAAGTTGGACT 909
 QY 751 TGGACTAGCAGGACCATGCTTCTGACCGGACGTGTACTCTTACAATACGCTGCTGTG 810
 DB 910 TGGACTAGCAGGACCATGCTTCTGACCGGACGTGTACTCTTACAATACGCTGCTGTG 969
 QY 811 CTGCGGCGCTGTGTGTGGCTGCAACTCTGCTGCGGTGTTGTTTCTGCTGTAGAG 870
 DB 970 CTGCGGCGCTGTGTGTGGCTGCAACTCTGCTGCGGTGTTGTTTCTGCTGTAGAG 1029
 QY 871 AAAAAGAGGATTTGCTATTCAATTTCAAAGAAATCTGAAAAGAGAGAGCAAAACAAGA 930
 DB 1030 AAAAAGAGGATTTGCTATTCAATTTCAAAGAAATCTGAAAAGAGAGAGCAAAACAAGA 1089
 QY 931 AACTGACAGAGAAAGTGAATGAAATCGGGTACAATTCAGATGAAACAAAAGACAC 990
 DB 1090 AACTGACAGAGAAAGTGAATGAAATCGGGTACAATTCAGATGAAACAAAAGACAC 1149
 QY 991 AGACACCGCTTCTCTCCCTCCAAATCCTGTGAATCCAGTGATCTCTGAAACAAAAGACAG 1050
 DB 1150 AGAACCGCTTCTCTCCCTCCAAATCCTGTGAATCCAGTGATCTCTGAAACAAAAGACAG 1209
 QY 1051 TAGCTGTGGCCCTCCTCACCAGCGGCTGATCAACGTCACCCAGCCAGCCAGCAAGTCATCC 1110
 DB 1210 TAGCTGTGGCCCTCCTCACCAGCGGCTGATCAACGTCACCCAGCCAGCCAGCAAGTCATCC 1269
 QY 1111 ACAGGCTTCTTTAAATCTGGCCAGTCTGAGAGAGTCAAGTAAATCAACTGTAGTATAG 1168
 DB 1270 ACAGGCTTCTTTAAATCTGGCCAGTCTGAGAGAGTCAAGTAAATCAACTGTAGTATAG 1327

RESULT 6
 AX380402
 LOCUS
 DEFINITION Sequence 7 from Patent WO0200710.
 ACCESSION AX380402
 VERSION AX380402.1 GI:19575332
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1
 Welch, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
 B7-like molecules and uses thereof
 Patent: WO 0200710-A 7 03-JAN-2002;
 Amgen, Inc. (US)
 FEATURES
 Location/Qualifiers
 1..1139
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 1..1134
 /note="unnamed protein product"
 /codon_start=1
 CDS

/protein_id="CAD28219.1"	
/db_xref="GI:19575333"	
/db_xref="REMTREMBL:CAD28219"	
translation="MVAGAMENRPPGSGSNEVIEGPNARVLKGSQARENCTVSGQ	
WKLIMWALSDMVLVSVRMEPIITNDRFTSQYDQGNLTSEMIIHNVPDSDGNIRC	
SLONSLRHSAYLVTVQVMBELFIPSVNLVVAENPEFVTCIPSHWTRLPDIISWELGLL	
VSHSYFVVPSPDLQSAVSILALTPOSNGLTICVATWKSILKARKSAVNTVTRCPQ	
DTGGINIPGVLISSLPISLPSLPTWGVGLGLAGTLMILTPTCTLIRCCCKRRCCGC	
NCCRCPCFCCRNKLNKRQTKKLRQVEMKPTATQMNKSPQTPLLSLPNPVPV	
ILNKETVAVALLTSLGILNVHPGQVHRLLLILWPVLRKSVQL"	
ORIGIN	
Query Match	66.0%; Score 771; DB 6; Length 1139;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 871; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	8 ATGGTGGCAGGACCATGAAATAGAGCCACCCCGGTTCTGGGTTCTGGTAAATGAAGTC 67
DB	1 ATGGTGGCAGGACCATGAAATAGAGCCACCCCGGTTCTGGGTTCTGGTAAATGAAGTC 60
QY	68 ATAGAAGGCCCCCAAAATGACAGAGTCCTGAAGGCTCCAGGCTCGCTTCAACTGCACC 127
DB	61 ATAGAAGGCCCCCAAAATGACAGAGTCCTGAAGGCTCCAGGCTCGCTTCAACTGCACC 120
QY	128 GTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTCAGTGACATGTGTGTGCTTAAGCGTC 187
DB	121 GTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTCAGTGACATGTGTGTGCTTAAGCGTC 180
QY	188 AGGCCCATGGAGCCCATCATCAATGACCGGTTCCACCTCTCAGAGGTACGACCCAGGGC 247
DB	181 AGGCCCATGGAGCCCATCATCAATGACCGGTTCCACCTCTCAGAGGTACGACCCAGGGC 240
QY	248 GGAACCTTACCTCGGAGATGATCATCCAAATGTGGAGCCCGAGTATCGGGGAAACATC 307
DB	241 GGAACCTTACCTCGGAGATGATCATCCAAATGTGGAGCCCGAGTATCGGGGAAACATC 300
QY	308 AGATGAGCGCTCCAGAACCTGCGCTCATGGATCTGTTACCTTACCTTACCGTCCAAAGTTATG 367
DB	301 AGATGAGCGCTCCAGAACCTGCGCTCATGGATCTGTTACCTTACCTTACCGTCCAAAGTTATG 360
QY	368 GGAGAGCTGTTCAATCCAGTGTTAAATCTTGTAGTCGCTGAGAAAGAACCTTGTGAAGTT 427
DB	361 GGAGAGCTGTTCAATCCAGTGTTAAATCTTGTAGTCGCTGAGAAAGAACCTTGTGAAGTT 420
QY	428 ACTTGTCTACCTCAGCTGCTGCTCCGATATTTCTCGGAGCTCGTCTCTG 487
DB	421 ACTTGTCTACCTCAGCTGCTGCTCCGATATTTCTCGGAGCTCGTCTCTG 480
QY	488 GTGAGCCATTCAAGCTATTATTTGTTCCGGAGCCAGCGACCTTCAAAGTSCAGTGAGC 547
DB	481 GTGAGCCATTCAAGCTATTATTTGTTCCGGAGCCAGCGACCTTCAAAGTSCAGTGAGC 540
QY	548 ATCTGTGCTGTACCCCAAGAGCAATGGGACTTTGATCTTGGTGGCTTACCTGGAAGAGC 607
DB	541 ATCTGTGCTGTACCCCAAGAGCAATGGGACTTTGATCTTGGTGGCTTACCTGGAAGAGC 600
QY	608 CTGAAGGCCCGCAAGTCTGCACTTAATCTCACTGTGATTCGGTGTCGCCCAAGACACT 667
DB	601 CTGAAGGCCCGCAAGTCTGCACTTAATCTCACTGTGATTCGGTGTCGCCCAAGACACT 660
QY	668 GGAGTGTGTTAATATTTCCAGGTGATTAATCAAGTTTACCGAGTTAGGTTTTTCATTG 727
DB	661 GGAGTGTGTTAATATTTCCAGGTGATTAATCAAGTTTACCGAGTTAGGTTTTTCATTG 720
QY	728 CCTACTTGGGGCAAAAGTTGACTTGACATGAGCAGGACCATGCTTCTGAGCCGACGTGT 787
DB	721 CCTACTTGGGGCAAAAGTTGACTTGACATGAGCAGGACCATGCTTCTGAGCCGACGTGT 780
QY	788 ACTCTTACAATACGCTGCTGCTGCCCGCTGTTGTTGGCTGCAACTGCTGCTGC 847
DB	781 ACTCTTACAATACGCTGCTGCTGCCCGCTGTTGTTGGCTGCAACTGCTGCTGC 840
QY	848 CGTTGTTGTTCTGCTGTAGAGAAAAAGAGGA 880

DB	841	CGTTGTTGTTCTGCTGTAGAGAAAAAGAGGA	873
RESULT 7			
AF121782		142742 bp	DNA linear PRI 02-FEB-1999
LOCUS			
DEFINITION		Homo sapiens chromosome 21q22.3 PAC 206A10, complete sequence.	
ACCESSION		AF121782	
VERSION		AF121782.1	GI:4210991
KEYWORDS		HTG.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE		1 (bases 1 to 142742)	
AUTHORS		Taudien,S., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B., Schatttevoy,R., Weber,J., Schilling,M., Menzel,U., Yaspo,M.L. and Rosenthal,A.	
TITLE		Direct Submission	
JOURNAL		Submitted (21-JAN-1999) Genome Analysis, Institute for Molecular Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany	
FEATURES		Location/Qualifiers	
1..142742			
/organism="Homo sapiens"			
/mol_type="genomic DNA"			
/db_xref="taxon:9606"			
/chromosome="21"			
/map="21q22.3"			
/clone="PAC 206A10"			
3..118			
/evidence=not_experimental			
/rpt_family="MIR"			
complement(140..415)			
/evidence=not_experimental			
/rpt_family="MLT2E"			
complement(668..881)			
/evidence=not_experimental			
/rpt_family="MER58A"			
complement(1078..1162)			
/note="GRAIL, score = 86.000%, comment = excellent shadow"			
/evidence=not_experimental			
1114..1403			
/evidence=not_experimental			
/rpt_family="MLT1I"			
1481..1603			
/evidence=not_experimental			
/rpt_family="L2"			
1576..1710			
/note="M2EF, score = 66.1%"			
/evidence=not_experimental			
1576..1710			
/note="GRAIL, score = 99.000%, comment = excellent"			
/evidence=not_experimental			
complement(1693..1741)			
/evidence=not_experimental			
/rpt_family="MIR"			
complement(1780..2326)			
/evidence=not_experimental			
/rpt_family="MLTIE"			
complement(1834..1955)			
/note="GRAIL, score = 50.000%, comment = good shadow"			
/evidence=not_experimental			
2629..2831			
/evidence=not_experimental			
/rpt_family="AluSx"			
3012..3122			
/evidence=not_experimental			
/rpt_family="MLT1J"			
3123..3478			
/evidence=not_experimental			
/rpt_family="THE1B"			
3479..3618			
repeat_region			

```

/evidence=not_experimental
/rpt_family="MLT1J"
complement(1302..3540)
/note="GRAIL, score = 65.000%, comment = good shadow"
exon
4674..4781
/evidence=not_experimental
/note="GRAIL, score = 42.000%, comment = marginal"
exon
4712..4749
/note="Xpound exon prediction, score = 70% (0%)"
repeat_region
5385..6104
/evidence=not_experimental
/rpt_family="HERVL"
exon
6682..6751
/note="MZEF, score = 95.1%"
exon
6746..6751
/evidence=not_experimental
/note="GenScan, score = 1.83%, comment = Initial_exon 6 bp
frame: 1 phase: 0"
repeat_region
6901..7339
/evidence=not_experimental
/rpt_family="MLT1C"
repeat_region
7347..7457
/evidence=not_experimental
/rpt_family="LIR16C"
repeat_region
7561..7696
/evidence=not_experimental
/rpt_family="MIR"
repeat_region
7819..8220
/evidence=not_experimental
/rpt_family="MLT1J"
exon
8097..8119
/note="GRAIL, score = 45.000%, comment = marginal"
repeat_region
8738..9232
/evidence=not_experimental
/rpt_family="MLT1C"
exon
9273..9420
/note="GRAIL, score = 71.000%, comment = good"
exon
9290..9371
/note="GRAIL, score = 81.000%, comment = excellent shadow"
repeat_region
9966..10041
/evidence=not_experimental
/note="homology = 100.00%, score = 38, counts = 2"
/rpt_type=tandem
/rpt_unit="tcttgcacactctctgtgagaggtcggttatcgc"
repeat_region
10244..10664
/evidence=not_experimental
/rpt_family="MLT1E"
repeat_region
10665..11028
/evidence=not_experimental
/rpt_family="MLT1F"
repeat_region
11318..11417
/evidence=not_experimental
/rpt_family="LIMC/D"
repeat_region
11503..11818
/evidence=not_experimental
/rpt_family="AluSg"
exon
11525..11777
/note="MZEF, score = 79.4%"
repeat_region
12346..12827
/evidence=not_experimental
/rpt_family="MLT1C"
repeat_region
12902..13180
/evidence=not_experimental
/rpt_family="AluSg"
exon
13057..13161
complement(13057..13161)
/note="MZEF, score = 76.1%"

```

```

/evidence=not_experimental
complement(13226..13364)
/rpt_family="LIMC2"
complement(13613..13908)
/evidence=not_experimental
/rpt_family="AluSx"
complement(14317..14614)
/evidence=not_experimental
/rpt_family="LIR37A"
complement(14962..15036)
/note="MZEF, score = 93%"
exon
15255..15299
/evidence=not_experimental
/rpt_family="LIPa6"
complement(15469..15497)
/note="Xpound exon prediction, score = 62% (0%)"
repeat_region
16045..16768
/evidence=not_experimental
/rpt_family="MER21A"
complement(16555..16655)
/note="GRAIL, score = 54.000%, comment = good"
repeat_region
16823..16850
/evidence=not_experimental
/rpt_family="MER34"
complement(16851..17208)
/evidence=not_experimental
/rpt_family="THEIA"
repeat_region
17209..17632
/evidence=not_experimental
/rpt_family="MER34"
complement(17633..17969)
/evidence=not_experimental
/rpt_family="AluSg1"
complement(18004..18334)
/evidence=not_experimental
/rpt_family="LIR17"
complement(18335..18632)
/evidence=not_experimental
/rpt_family="HERV17"
exon
18352..18420
/note="GRAIL, score = 51.000%, comment = good shadow"
exon
18397..18594
/evidence=not_experimental
/note="GRAIL, score = 66.000%, comment = good"
repeat_region
18643..24647
/evidence=not_experimental
/rpt_family="HERV17"
complement(19228..19361)
/note="MZEF, score = 92.8%"
exon
20900..21135
/note="MZEF, score = 51.4%"
exon
21150..21275
/evidence=not_experimental
/note="Xpound exon prediction, score = 74% (0%)"

```

```

Query Match 27.3%; Score 319; DB 9; Length 142742;
Best Local Similarity 100.0%; Pred. No. 3.5e-176;
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 GGTTCCTGGGTCTGTAATGAAGTCATAGAGGCCCAAAATGCAAGAGTCTCTGAAGGC 103
|||||
Db 44274 GGTTCCTGGGTCTGTAATGAAGTCATAGAGGCCCAAAATGCAAGAGTCTCTGAAGGC 44333
|||||

QY 104 TCCCAAGGCTCGTTCAACTGCACCGCTCCCAAGGCTGCAAGCTCATCATGTGGGCTCTC 163
|||||
Db 44334 TCCCAAGGCTCGTTCAACTGCACCGCTCCCAAGGCTGCAAGCTCATCATGTGGGCTCTC 44393
|||||

```

Wed Sep 15 10:41:07 2004

164 AGTCACATGTTGGTCTAAGCGTCAGCGCCATGAGCCCATCATCACCAATGACGGCTTC 223
 44394 AGTCACATGTTGGTCTAAGCGTCAGCGCCATGAGCCCATCATCACCAATGACGGCTTC 44453
 224 ACCTCTCAGAGTACGACGCGCGGGAACCTTACCTCGGAGATGATCATCCACAATGTG 283
 44454 ACCTCTCAGAGTACGACGCGCGGGAACCTTACCTCGGAGATGATCATCCACAATGTG 44513
 284 GAGCCCACTGATTCGGGGAACATCATGATGACGCTCCAGACAGTCGCTGCTGATGATCT 343
 44514 GAGCCCACTGATTCGGGGAACATCATGATGACGCTCCAGACAGTCGCTGCTGATGATCT 44573
 344 GCTTACCTTACCTGCAAG 362
 44574 GCTTACCTTACCTGCAAG 44592

RESULT 8
 HS21C080 340000 bp DNA linear PRI 24-MAY-2000
 LOCUS Homo sapiens chromosome 21 segment HS21C080.
 DEFINITION AL163280 AP001735 BA000005
 ACCESSION AL163280.2 GI:7717369
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 340000)
 Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
 Park H.S., Toyoda A., Ishii K., Toki Y., Choi D.K., Soeda E.,
 Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
 Polley A., Menzel J., Delabar J., Kumpf K., Lehmann R.,
 Patterson D., Reichwald K., Rump A., Schilhabel M., Schudy A.,
 Zimmermann W., Rosenthal A., Kudoh J., Shibuya K., Kawasaki K.,
 Asakawa S., Shintani A., Sasaki T., Nagamine K., Mitsuoyama S.,
 Antonarakis S.E., Minoshima S., Shimizu N., Nordieck G.,
 Hornischer K., Brandt P., Scharfe M., Schoen O., Desario A.,
 Reichelt J., Kauer G., Bloeker H., Ranser J., Beck A., Klages S.,
 Hennig S., Rieselmann L., Dagand E., Wehrmeyer S., Borzym K.,
 Gardiner K., Nizetic D., Francis F., Lehrach H., Reinhardt R. and
 Yaspo M.L.

Direct Submission
 Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing
 Consortium: * RIKEN Genomic Sciences Center, Human Genome Research
 Group * Institute of Molecular Biotechnology, Genome Analysis *
 Keio University School of Medicine, Dept. of Molecular Biology *
 GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular
 Genetics (addresses see below)
 The Chromosome 21 Mapping and Sequencing Consortium consists of
 * RIKEN Genomic Sciences Center, Human Genome Research Group, *
 Sagamiyara 228-8555, Japan,
 * e-mail: sakaki@gsc.riken.go.jp
 * URL: http://hgp.gsc.riken.go.jp/
 and
 * Institute of Molecular Biotechnology, Genome Analysis, *
 Beutenbergstrasse 11, D-07745 Jena, Germany,
 * e-mail: gscj-submit@genome.imb-jena.de
 * URL: http://genome.imb-jena.de/
 and
 * Keio University School of Medicine, Dept. of Molecular Biology, *
 Tokyo 160-8582, Japan,
 * e-mail: shimizu@db.med.keio.ac.jp
 * URL: http://adenine.dmb.med.keio.ac.jp/
 and
 * GBF, Dept. of Genome Analysis,
 * Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e-mail:
 info.genome@gbf.de
 * URL: http://genome.gbf.de/
 and
 * Max-Planck Institute for Molecular Genetics,
 * Innestrasse 73, D-14195 Berlin, Germany,
 * e-mail: info-chr21@molgen.mpg.de

repeat_region
 repeat_region
 repeat_region
 repeat_region
 repeat_region
 repeat_region
 gene
 mRNA
 CDS

* URL: http://chr21.rz-berlin.mpg.de/.
 Location/Qualifiers
 1. 340000
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"
 <1.125946
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"
 /clone="P160L9, 5' partial"
 /clone_lib="RPC11,3-5 PAC library"
 /note="Accession No. AF121897"
 73410. 243533
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"
 /clone="P70124"
 /clone_lib="RPC11,3-5 PAC library"
 /note="Accession No. AF064860"
 224137. >340000
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"
 /clone="P206A10, 3' partial"
 /clone_lib="RPC11,3-5 PAC library"
 /note="Accession No. AF121782"
 286628. >340000
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"
 /clone="BAC-291B3, 3' partial"
 /clone_lib="BAC library"
 /note="Accession No. AF064857"
 1088. 1586
 /note="LIM4"
 /rpt_family="LINE/L1"
 /rpt_type=DISPERSED
 2398. 2495
 /note="(TCTA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 complement (2496. 2878)
 /note="THE1C"
 /rpt_family="LTR/MaLR"
 /rpt_type=DISPERSED
 2879. 2931
 /note="(TCTA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 2932. 3004
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3005. 3074
 /note="(TCCA)n"
 /rpt_family="Simple_repeat"
 /rpt_type=TANDEM
 3473. 19157
 /gene="SH3BGR"
 join(<3473. 3565,12649. 12678,15331. 15460,18739. 19157)
 /gene="SH3BGR"
 join(3473. 3565,12649. 12678,15331. 15426)

```

/ gene="SH3BGR"
/ note="Accession No. X93498"
/ codon_start=1
/ product="21-Glutamic Acid Rich protein 21-GARP"
/ protein_id="CAB90445.1"
/ db_xref="GI:7717370"
/ db_xref="GOA:P55822"
/ db_xref="SWISS-PROT:P55822"
/ translation="GSEKABEGTEAQKEGSEVGNLPEAQKNEEBEGTATERTEE"
/ name="SH3BGR"
/ number=4
/ repeat_region
/ note="AluSx"
/ rpt_family="SINE/Alu"
/ rpt_type=DISPERSED
/ complement(6887..7067)
/ note="L1MD3"
/ rpt_family="LINE/L1"
/ rpt_type=DISPERSED
/ complement(8482..8762)
/ note="AluJo"
/ rpt_family="SINE/Alu"
/ rpt_type=DISPERSED
/ complement(8865..8964)
/ note="L2"
/ rpt_family="LINE/L2"
/ rpt_type=DISPERSED
/ 9131..9424
/ note="AluSq"
/ rpt_family="SINE/Alu"
/ rpt_type=DISPERSED
/ complement(9723..9882)
/ note="L2"
/ rpt_family="LINE/L2"
/ rpt_type=DISPERSED
/ complement(9989..10077)
/ note="L2"
/ rpt_family="LINE/L2"
/ rpt_type=DISPERSED
/ complement(10078..10429)
/ note="THE1B"
/ rpt_family="LTR/MaLR"
/ rpt_type=DISPERSED
/ complement(10430..10651)
/ note="L2"
/ rpt_family="LINE/L2"
/ rpt_type=DISPERSED
/ 10980..11000
/ note="(TTTA)n"
/ rpt_family="Simple_repeat"
/ rpt_type=TANDEM
/ complement(11002..11282)
/ note="AluSp"
/ rpt_family="SINE/Alu"
/ rpt_type=DISPERSED
/ complement(11315..11556)
/ note="AluJo"
/ rpt_family="SINE/Alu"
/ rpt_type=DISPERSED
/ 11702..11753
/ note="MIR"
/ rpt_family="SINE/MIR"
/ rpt_type=DISPERSED
/ 12649..12678
/ gene="SH3BGR"
/ number=5
/ repeat_region
/ note="AluSx"
/ rpt_family="SINE/Alu"
/ rpt_type=DISPERSED
/ complement(13598..13892)

```

```

Query Match      27.3%; Score 319; DB 9; Length 340000;
Best Local Similarity 100.0%; Pred. No. 3.6e-176;
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 GGTTCCTGGGCTCTGTAATCAAGTCATAGAAGCGCCCAAAATGCAAGAGTCTCTGAAGGGC 103
      |||
Db 268454 GTTTCCTGGGCTCTGTAATCAAGTCATAGAAGCGCCCAAAATGCAAGAGTCTCTGAAGGGC 268513
      |||

QY 104 TCCCAGGCTCGCTTCAACTGCAACCGTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTC 163
      |||
Db 268514 TCCCAGGCTCGCTTCAACTGCAACCGTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTC 268573
      |||

QY 164 AGTGACATGTGGTCTTAAGCGTCAAGCGCTAGGCCATGAGGCCCATCATACCAATGACCGCTTC 223
      |||
Db 268574 AGTGACATGTGGTCTTAAGCGTCAAGCGCTAGGCCATGAGGCCCATCATACCAATGACCGCTTC 268633
      |||

QY 224 ACCTCTCAGAGGTACGACCGGGGGAACTTCACTCGGAGATGATCATCCCAATGTG 283
      |||
Db 268634 ACCTCTCAGAGGTACGACCGGGGGAACTTCACTCGGAGATGATCATCCCAATGTG 268693
      |||

QY 284 GAGCCAGTGATTCGGGGAACATCAGATGCAAGCTCCAGAACAGTCGGCTCATGGATCT 343
      |||
Db 268694 GAGCCAGTGATTCGGGGAACATCAGATGCAAGCTCCAGAACAGTCGGCTCATGGATCT 268753
      |||

QY 344 GCTTACCTTACCGTCCAAG 362
      |||
Db 268754 GCTTACCTTACCGTCCAAG 268772
      |||

RESULT 9
CH179K04      182532 bp DNA linear PRI 16-OCT-2003
LOCUS Pan troglodytes chromosome 22 BAC CH251-179K04, complete sequence.
DEFINITION AL954228
ACCESSION AL954228
VERSION AL954228.1 GI:37619870
KEYWORDS HTG.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 182532)
.
The Chimpanzee Chromosome 22 Sequencing Consortium
Chimpanzee chromosome 22 genomic sequence
Unpublished
2 (bases 1 to 182532)
Schafte, M., Berg, C., Conrad, A., Hornischer, K., Loehnert, T.H.,
Ludwig, M., Thies, S., Weber, K. and Bloeker, H.
Direct Submission
Submitted (03-JUN-2003) GBF, Dept. of Genome Analysis, Mascheroder
Weg 1, D-38124 Braunschweig, Germany, E-mail: info.genome@gbf.de
The Chimpanzee Chromosome 22 Sequencing Consortium consists of :
Shanghai, China
Taiwan:
----- Genome Center
Center: GBF, Braunschweig
Center code: GBF
Web site: http://genome.gbf.de/
Contact: info.genome@gbf.de
----- Project Information
Center project name:
Center clone name: CH251-179K04
----- Summary Statistics
Sequencing vector: ###;
Chemistry: Dye-terminator-amersham: ## of reads
Chemistry: Dye-primer-amersham: ## of reads
Assembly program: ##
Consensus quality: 0 bases at least Q40
Consensus quality: 0 bases at least Q30
Consensus quality: 0 bases at least Q20
Estimated insert size: ##; agarose-fp estimation
Estimated insert size: 182532; sum-of-contigs estimation

```

```

PROGRAMS AND PARAMETERS USED FOR ANNOTATION:
+ Analysis and annotation were performed with the automatic
+ 'first-pass' annotation and submission tool
+ 'AnnoMitter' (Hornischer & Bloecker)
+ Programs used by 'AnnoMitter':
+
+***** Location/Qualifiers
+*****
+***** 1. 182532
+***** /organism="Pan troglodytes"
+***** /mol_type="genomic DNA"
+***** /db_xref="taxon:9598"
+***** /chromosome="22"
+***** /clone="CH251-179K04"
+*****
+***** misc_feature
+*****
+***** 1. 182532
+***** /note="assembly fragment-clone end:T7-vector side:left
+***** assembly_fragment-clone end:SP6-vector side:right"
+*****
+***** ORIGIN
+*****
Query Match 20.6%; Score 241; DB 9; Length 182532;
Best Local Similarity 99.7%; Pred. No. 5.1e-130;
Matches 291; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 361 AGTTATGGAGAGCTCTTCAATCCAGTGTAACTTGTAGTCGCTGAGAAATGAACCTTG 420
DB 17792 AGTTATGGAGAGCTCTTCAATCCAGTGTAACTTGTAGTCGCTGAGAAATGAACCTTG 17851

QY 421 TGAAGTACTTGTCTACCTCCACTGGAGCTGGCTCCGGATATTTCTTGGAGCTGG 480
DB 17852 TGAAGTACTTGTCTAGCTCAGCTGGAGCTGGCTCCGGATATTTCTTGGAGCTGG 17911

QY 481 TCTCTCTGGTCAGCAATCAAGCTATTTATTTTGTTCGGAGCCAGCACTTCAAGTGC 540
DB 17912 TCTCTCTGGTCAGCAATCAAGCTATTTATTTTGTTCGGAGCCAGCACTTCAAGTGC 17971

QY 541 AGTGAGCACTCTGGCTGACCCACAGAGCAATGGGACTTTCAGTTCGCTGGCTACCTG 600
DB 17972 AGTGAGCACTCTGGCTGACCCACAGAGCAATGGGACTTTCAGTTCGCTGGCTACCTG 18031

QY 601 GAAGAGCTCAAGCCCGCAAGCTGCAACTGCTCAACTGTAATCTCACTGTGATCGG 652
DB 18032 GAAGAGCTCAAGCCCGCAAGCTGCAACTGCTCAACTGTAATCTCACTGTGATCGG 18083

RESULT 10
LOCUS BS000162 156288 bp DNA linear PRI 07-OCT-2003
DEFINITION Pan troglodytes chromosome 22 clone:PTB-060F12, map 22, complete
sequences.
ACCESSION BS000162
VERSION BS000162.1 GI:37537429
KEYWORDS HTG.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.

REFERENCE 1 The Chimpanzee Chromosome 22 Sequencing Consortium.
AUTHORS DNA sequence of chimpanzee chromosome 22 and its evolutionary
TITLE implications
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 156288)
AUTHORS Saitou,N., Kim,C., Kitano,T., Oota,S., Shimada,M., Kryukov,K.,
TITLE Tomiki,T. and Kohara,Y.
JOURNAL Submitted (15-MAY-2003) Naruya Saitou, National Institute of
Genetics (NIG), Division of Population Genetics, 1111 Yata,
Mishima, Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
URL:http://sayer.lab.nig.ac.jp/, Tel:81-55-981-6796,
Fax:81-55-981-6789)
COMMENT The Chimpanzee Chromosome 22 Sequencing Consortium consists of:
*Chinese National Human Genome Center at Shanghai, China;

```

```

*GBF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
of Molecular Biotechnology, Jena, Germany; *KRIIBB Genome Research
Center, Daejeon, Korea;
*Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
*National Institute of Genetics, Mishima, Japan;
*National Yang Ming University Genome Research Center, Taipei,
Taiwan;
*RIKEN Genomic Sciences Center, Yokohama, Japan.
----- Genome Center
Center: National Institute of Genetics
Center code: NIG
Web site: http://sayer.lab.nig.ac.jp/
Contact: nsaitou@genes.nig.ac.jp
----- Project Information
Center project name: The Chimpanzee Chromosome 22 Sequencing Project
Center clone name: PTB-060F12
----- Summary Statistics
Sequencing vector: pUC118; 100% of reads
Chemistry: Dye-terminator Big Dye and dGTP; 100% of reads Assembly
Program: Phrap; version 0.990329
Consensus quality: 156,231 bases at least Q40
Consensus quality: 156,288 bases at least Q30
Consensus quality: 156,288 bases at least Q20
-----
This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30);
an attempt was made to resolve all sequencing problems, such as
compressions and repeats; all regions were covered by at one
plasmid
subclone or more than one M13 subclone;
-----
Source information:
The PTB1 chimpanzee BAC library was prepared from DNA isolated from
cultured cells established from the blood of a single male
chimpanzee.
Clones may be obtained from Asao Fujiyama and co-workers
(http://www.gsc.riken.go.jp).
VECTOR: pKS145
-----
Sequence Quality Assessment:
This entry has been annotated with sequence
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in
10,000 bp.
-----
Neighboring clones: CH251-179K04(left) and PTB-103H04(right).
-----
FEATURES
source
1. 156288
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="22"
/clone="PTB-060F12"
/clone_lib="PTB1 chimpanzee BAC"
-----
ORIGIN
Query Match 12.4%; Score 145; DB 9; Length 156288;
Best Local Similarity 100.0%; Pred. No. 3.4e-73;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 473 GAGCTCGGTCTCTCTGGTCAGCAATCAAGCTATTTATTTTTCGGAGCCAGCACTT 532
DB 1 GAGCTCGGTCTCTCTGGTCAGCAATCAAGCTATTTATTTTTCGGAGCCAGCACTT 60

QY 533 CAAAGTGAGTGAGCATCTCTGCTGACCCACAGAGCAATGGGACTTTGACTTGCCTG 592
DB 61 CAAAGTGAGTGAGCATCTCTGCTGACCCACAGAGCAATGGGACTTTGACTTGCCTG 120

QY 593 GCTACTGGAAGAGCTGAAGGCC 617

```

```

Db          121 GCTACTGGAGAGCCTGAGGCC 145

RESULT 11
AF064857    199665 bp      DNA      linear      PRI 23-JUN-1998
LOCUS      Homo sapiens BAC derived from chromosome 21q22.3, complete
DEFINITION sequence, containing PEB19 (PCP4) gene.
ACCESSION  AF064857
VERSION     AF064857.1  GI:3171149
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 199665)
AUTHORS   Taudien,S., Nordsiek,G., Korenberg,J., Drescher,B., Weber,J.,
          Schattevoy,R. and Rosenthal,A.
TITLE     Direct Submission
JOURNAL   Submitted (11-MAY-1998) Genome Analysis, Institute for Molecular
          Biotechnology, Heutenbergstrasse 11, Jena 07745, Germany
FEATURES   source
            1..199665
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /chromosome="21"
             /map="21q22.3"
             complement(1..383)
             /evidence=not_experimental
             /rpt_family="L1"
             747..814
             /evidence=not_experimental
             /rpt_family="MIR"
             complement(1345..1500)
             /evidence=not_experimental
             /rpt_family="L1PA4"
             1499..1588
             /evidence=not_experimental
             /rpt_family="L1PA2"
             complement(1584..2179)
             /evidence=not_experimental
             /rpt_family="L1PA2"
             2174..3405
             /evidence=not_experimental
             /rpt_family="L1PA2"
             3388..3454
             /evidence=not_experimental
             /rpt_family="L1p"
             3968..4086
             /evidence=not_experimental
             /rpt_family="MIR"
             complement(6245..6668)
             /evidence=not_experimental
             /rpt_family="L1ME3A"
             6671..6969
             /evidence=not_experimental
             /rpt_family="AluJb"
             complement(7057..7376)
             /evidence=not_experimental
             /rpt_family="MER2"
             complement(7387..7491)
             /evidence=not_experimental
             /rpt_family="L1ME3A"
             7781..8020
             /evidence=not_experimental
             /rpt_family="MER58A"
             8650..8823
             /evidence=not_experimental
             /rpt_family="L1PA5"
             complement(10456..10720)
             /evidence=not_experimental
             /rpt_family="L2"

repeat_region complement(13514..13814)
/evidence=not_experimental
/rpt_family="AluSx"
repeat_region complement(13823..14116)
/evidence=not_experimental
/rpt_family="AluJo"
repeat_region complement(15662..15785)
/evidence=not_experimental
/rpt_family="MER5A"
repeat_region complement(18994..19347)
/evidence=not_experimental
/rpt_family="THE1B"
repeat_region 19378..19470
/note="homology = 81.70%, score = 28, counts = 3"
/evidence=not_experimental
/rpt_type=tandem
repeat_region /rpt_unit="agtatctattctgattgggcagtgctctac"
complement(20371..20556)
/evidence=not_experimental
/rpt_family="MER3"
repeat_region 21213..21500
/evidence=not_experimental
/rpt_family="AluSx"
repeat_region 22291..22567
/evidence=not_experimental
/rpt_family="MER39b"
repeat_region complement(22583..23201)
/evidence=not_experimental
/rpt_family="L1PA2"
repeat_region 23202..23471
/evidence=not_experimental
/rpt_family="MER39b"
repeat_region 24223..24454
/evidence=not_experimental
/rpt_family="MSTB-internal"
repeat_region 24455..24889
/evidence=not_experimental
/rpt_family="MSTB"
repeat_region complement(25434..25767)
/evidence=not_experimental
/rpt_family="MER58B"
repeat_region 25993..26294
/evidence=not_experimental
/rpt_family="AluSq"
repeat_region 27222..27283
/evidence=not_experimental
/rpt_family="MIR"
repeat_region complement(28932..29335)
/evidence=not_experimental
/rpt_family="MSTA"
repeat_region complement(29462..30001)
/evidence=not_experimental
/rpt_family="MER34"
repeat_region complement(30308..30676)
/evidence=not_experimental
/rpt_family="MLT1B"
repeat_region 31012..31364
/evidence=not_experimental
/rpt_family="MLT1A2"
repeat_region complement(32507..32738)
/evidence=not_experimental
/rpt_family="MSTd"
repeat_region complement(32766..32980)
/evidence=not_experimental
/rpt_family="MSTd"
repeat_region 33468..33678
/evidence=not_experimental
/rpt_family="THE1B"
repeat_region 34563..34620
/note="homology = 93.10%, score = 21, counts = 2"
/evidence=not_experimental
/rpt_type=tandem
/rpt_unit="ttccagaaagaatacgtactttaggca"

```

repeat_region	complement(35785..35824) /evidence=not_experimental /rpt_family="MLT2CB"	
repeat_region	complement(35924..36345) /evidence=not_experimental /rpt_family="MLT2CB"	
repeat_region	complement(36403..36475) /evidence=not_experimental /rpt_family="MLT2CB"	
repeat_region	complement(36717..36828) /evidence=not_experimental /rpt_family="MER67C"	
repeat_region	complement(37064..37364) /evidence=not_experimental /rpt_family="AluY"	
repeat_region	37440..37501 /note="homology = 98.40%, score = 29, counts = 2" /evidence=not_experimental /rpt_type="tandem"	
repeat_region	37518..38158 /evidence=not_experimental /rpt_family="L2"	
repeat_region	38477..39376 /evidence=not_experimental /rpt_family="L2"	
repeat_region	39552..39861 /evidence=not_experimental /rpt_family="MLT1D"	
repeat_region	39884..40122 /evidence=not_experimental /rpt_family="MLT1D"	
repeat_region	complement(40243..40542) /evidence=not_experimental /rpt_family="AluSx"	
repeat_region	40543..41180 /evidence=not_experimental /rpt_family="L2"	
repeat_region	complement(42066..42114) /evidence=not_experimental /rpt_family="MIR"	
repeat_region	42138..42317 /evidence=not_experimental /rpt_family="L2"	
repeat_region	complement(42332..42710) /evidence=not_experimental /rpt_family="L1PB1"	
repeat_region	complement(42758..46751) /evidence=not_experimental /rpt_family="L1PB1"	
repeat_region	complement(46792..47084) /evidence=not_experimental /rpt_family="L1MA"	
repeat_region	47189..47992 /evidence=not_experimental /rpt_family="L2"	
repeat_region	complement(48002..48132) /evidence=not_experimental /rpt_family="MIR"	

Query Match	8.6%; Score 100; DB 9; Length 199665;
Best Local Similarity	100.0%; Pred. No. 1.5e-46;
Matches 100; Conservative	0; Mismatches 0; Indels 0; Gaps 0;

QY	1069	CCAGCGGGTGATCAACCTCCACCCAGGCGAGTATCCACAGGCTCTTTTAATCT	1128
Db	17555	CCAGCGGGTGATCAACCTCCACCCAGGCGAGTATCCACAGGCTCTTTTAATCT	17614

QY	1129	GGCCAGTCTTGAGAGGTCAAGTAATACACTAGTATAG	1168
Db	17615	GGCCAGTCTTGAGAGGTCAAGTAATACACTAGTATAG	17654

RESULT 12	RP43002119	192219 bp	DNA	linear	PRI 19-NOV-2003
LOCUS	Pan troglodytes chromosome 22 BAC RP43-002119, complete sequence.				
DEFINITION	AL954227				
ACCESSION	AL954227.3	GI:38453654			
VERSION	HTG.				
KEYWORDS	Pan troglodytes (chimpanzee)				
SOURCE	Pan troglodytes				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.				
REFERENCE	1 (bases 1 to 192219)				
AUTHORS	The Chimpanzee Chromosome 22 Sequencing Consortium.				
TITLE	Chimpanzee chromosome 22 genomic sequence				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 192219)				
AUTHORS	Scharfe, M., Berg, C., Conrad, A., Hornischer, K., Loehnert, T.H., Ludwig, M., Thies, S., Weber, K. and Bloecker, H.				
TITLE	Direct Submission (25-JUN-2003) GBF, Dept. of Genome Analysis, Mascheroder Weg 1, D-38124 Braunschweig, Germany, E-mail: info.genome@gbf.de				
JOURNAL	On Nov 19, 2003 this sequence version replaced gi:38228900.				
COMMENT	The Chimpanzee Chromosome 22 Sequencing Consortium consists of : *Chinese National Human Genome Center at Shanghai, Shanghai, China *GBF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute of Molecular Biotechnology, Jena, Germany; *KRIBB Genome Research Center, Daejeon, Korea; *Max-Planck Institute for Molecular Genetics, Berlin, Germany; *National Institute of Genetics, Mishima, Japan; *National Yang Ming University Genome Research Center, Taipei, Taiwan; *RIKEN Genomic Sciences Center, Yokohama, Japan. PROGRAMS AND PARAMETERS USED FOR ANNOTATION: ***** + Analysis and annotation were performed with the automatic + 'first-pass' annotation and submission tool + 'AnnoMitter' (Hornischer & Bloecker). + Programs used by 'AnnoMitter': ***** ----- Genome Center Center: GBF, Braunschweig Center code: GBF Web site: http://genome.gbf.de/ Contact: info.genome@gbf.de ----- Project Information Center project name: Center clone name: RP43-002119 ----- Summary Statistics Sequencing vector: ##, Chemistry: Dye-terminator-amersham: ## of reads Chemistry: Dye-primer-amersham: ## of reads Assembly program: ## Consensus quality: 0 bases at least Q40 Consensus quality: 0 bases at least Q30 Consensus quality: 0 bases at least Q20 Estimated insert size: ##; agarose-fp estimation Estimated insert size: 192219; sum-of-contigs estimation *****				

FEATURES	Location/Qualifiers
source	1..192219
	/organism="Pan troglodytes"
	/mol_type="genomic DNA"
	/db_xref="taxon:9598"
	/chromosome="22"
	/clone="RP43-002119"
misc_feature	1..192219
	/note="assembly_fragment-clone end:T7-vector side:left assembly_fragment-clone end:SP6-vector side:right"

ORIGIN	
Query Match	6.2%; Score 72; DB 9; Length 192219;

```

Best Local Similarity 100.0%; Pred. No. 5.5e-30;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 GGGAACTTCACCTCGAGATGATCATCCACATATGGAGCCAGGATTCGGGGAAACATC 307
      |||
Db 191622 GGGAACTTCACCTCGAGATGATCATCCACATATGGAGCCAGGATTCGGGGAAACATC 191681

QY 308 AGATCAGCCTC 319
      |||
Db 191682 AGATCAGCCTC 191693

RESULT 13
LOCUS AF064860 170121 bp DNA linear PRI 05-MAR-2002
DEFINITION Homo sapiens chromosome 21 clone PAC 70124 map 21q22.3, complete
sequence.
ACCESSION AF064860
VERSION AF064860.2 GI:18958624
KEYWORDS HTG; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 170121)
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
Park,H.-S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.-K., Soeda,E.,
Okii,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
Patterson,D., Reichwald,K., Rump,A., Schillhabel,M.B., Schudy,A.,
Zimmermann,W., Rosenthal,A., Kudon,J., Kawasaki,K., Asakawa,S.,
Shintani,A., Saeki,T., Nagamine,K., Mitsuyama,S.,
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,
Hornischer,K., Brandt,P., Scharfe,M., Schoen,O., Desario,A.,
Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S.,
Hennig,S., Rieselmann,L., Dagand,E., Haaf,T., Wehrmeyer,S.,
Borzym,K., Gardiner,K., Nizetic,D., Francis,F., Lehrach,H.,
Reinhardt,R. and Yaspo,M.Laure.
The DNA sequence of human chromosome 21
Nature 405 (6784), 311-319 (2000)
REFERENCE 2 10830953
Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.
Direct Submission
Submitted (12-MAY-1998) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
REFERENCE 3 (bases 1 to 170121)
Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L., Rosenthal,A., Yaspo,M.-L. and
Rosenthal,A.
Direct Submission
Submitted (27-FEB-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
REFERENCE 4 (bases 1 to 170121)
Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.
Direct Submission
Submitted (05-MAR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
COMMENT On Feb 27, 2002 this sequence version replaced gi:3171153.
FEATURES
Location/Qualifiers
1..170121
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="21"
/map="21q22.3"
/clone="PAC 70124"

ORIGIN
Query Match 3.4%; Score 40; DB 9; Length 170121;

```

```

Best Local Similarity 100.0%; Pred. No. 4.8e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CATGTGGCAGAGCCATGGAAATAGAGACCCACCCGCT 46
      |||
Db 97137 CATGTGGCAGAGCCATGGAAATAGAGACCCACCCGCT 97176

RESULT 14
LOCUS AB090820 6470 bp DNA linear INV 25-MAR-2003
DEFINITION Anopheles gambiae retrotransposon R7Ag1 DNA, complete sequence.
ACCESSION AB090820
VERSION AB090820.1 GI:28569877
KEYWORDS Anopheles gambiae (African malaria mosquito)
SOURCE ORGANISM
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.
REFERENCE 1
Kojima,K.K. and Fujiwara,H.
Evolution of Target Specificity in R1 Clade Non-LTR
Retrotransposons
Mol. Biol. Evol. 20 (3), 351-361 (2003)
JOURNAL 22531580
MEDLINE 12644555
PUBMED
REFERENCE 2 (bases 1 to 6470)
Kojima,K.K. and Fujiwara,H.
Direct Submission
Submitted (26-AUG-2002) Kenji K Kojima, University of Tokyo,
Department of Integrated Biosciences, Graduate School of Frontier
Sciences, Bioscience Building 501, 5-1-5 Kashiwanoha, Kashiwa,
Chiba 277-8562, Japan (E-mail:kk27513@mail.ecc.u-tokyo.ac.jp,
Tel:81-4-7136-3661, Fax:81-4-7136-3660)
FEATURES
Location/Qualifiers
1..6470
/organism="Anopheles gambiae"
/mol_type="genomic DNA"
/db_xref="taxon:7165"
repeat_region 1..6470
/notes="non-LTR retrotransposon"
/transposon="R7Ag1"
959..2542
/notes="ORF1"
/codon_start=1
/product="gag-like protein"
/protein_id="BAC57915.1"
/db_xref="GI:28569878"
/translation="MDKQLRGRTISYDERPAVIRKLGSEKKLGTIVEPSSAGVPAR
WTATGVKSGATATKLTATSTPSTGTVRRMLADAKADNETTGVIVKRLDEIQILRLQ
MEASNLQEAQREAREAREADARVREAREHEERKEKELFNALLAGTLGTSGLRES
QQLQREQLLRMMESQQRQEQQLQDQQRQWRQOQQRQOQRLPAQOWPTVOQS
VRAQGVTVTSASSAVPDAGTVGVVGNQGNKONGVNLPOQSAQROPAPRHQOQW
PHQOQGOQOQGHQIQEKRRPRKRPDEIVVVPAGVSPKEMVYKINTPRIADPQ
RQIGVRRTPRDHLLPLPSRDSALKDIIQEVIGERSVTVTMAEAVVLDGDNM
IDEEAIKALMTTLGKSLVATVNLWERMEDMTKRARVLRPAEAPVAKVORRLGLTY
LCGHAHRIGAAACGGQPSN"
2544..6212
/notes="ORF2"
/codon_start=1
/product="reverse transcriptase"
/protein_id="BAC57916.1"
/db_xref="GI:28569879"
/translation="MEVLQINVNRSRSDIALNTMRVERADVCLMWELHVSRRNGN
WVADRLGKVAIIASSETVPQOVSVTQSGIAAARINDVLFICCVSPAGVSEPEV
MQRIDVLARGHPRVVFAGDNLAWHTANGSCRINAKGAAVQVLVSLGVLNLTGAPT
FLNGVAPSVVDVAFASSIAGVNTIPEHRWRIISYISNVHVIYRFVAGELLQRP
ADSRQEGPSTRESGTRWRTRFDALFGLDVAFTFRTVSASLERVMTACDAA
MARVFPQSHSGRPAYWTPATVLCNCLAKLEAAIDEEIOIAASDILQVRTA
LDSSITTSKKEHFDILGLADETGQWYRVLSPLSGSWTAREDSVLEIVITLFP
PQHPVDWTPASPGQVLERGEPEPVDVNEQELLDIASSLNPKAPGLQGVNAAUTAA

```

IRKHTDIFKLFQBLCDNERPDEWKKOKLALIPKPKPPGLASSFRPILILNPGKV
YERLLSRINDVIDEPSPRIENQYGERGRSTVOAIOVLVDAGSHAMSGRTNRRD
KRCLLVVALDVRNAPATNSQCIATLEDKGVPRQLRNILRDYFANRELIVVDADGPV
TRRTVAGVQOSILGPTLWNIIMYDGLVRELPEGASVIGYADDIVVMARGCTPQBAAL
VAEQADIAAWMEDHILQLAPEKTEGVMISLRRGQKVPFRVGDTHHSKQSIIRYL
GVQIHDLHSVPHVLSAKALRVVGVTVAVNRHSGPQAKRLLAAVAESIIRYAA
PWSBATTLOWCKLOVORPLARGVTSFVSVAJETGVAGLGVPPRLIVREDARC
HRRLLAAGASRKDIJLEROGTFOEMORANDAAAAPTASRYAVAHMIPDLHLWM
SRHGEVDFHLSQVLTGHGYFRYHLVHGFAPASACPCPGSVESVAHVLPQCEVFEH
IRVELLGYTSDPVNENMLKLESPEWRNMSIQBAARKITKVLQQLWREDELQNLQ
AHLAALPRAAADVADPDEGVSDGVGVAELFRNGRARRTRRRRAAREEVEVLLA
SAMAAERELRDSILMAAARAEAGEAPPPIPMRRRGLPPSPRTVRAHERLXLQRL
YQORAREGTLPTVHGRNRRSRAPSADTIRRRMRREMERLRTARRVPSNOQVRE
VLSADALAAITAITSGR"

ORIGIN

Query Match 2.1%; Score 25; DB 3; Length 6470;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 840 GCTGCTGCCGTTGTTCTTCTGCTG 864
|||||
Db 1574 GCTGCTGCCGTTGTTCTTCTGCTG 1550

RESULT 15

AY119603 5178 bp mRNA linear INV 15-JUN-2002
Drosophila melanogaster LD27161 full insert cDNA.
ACCESSION AY119603
VERSION AY119603.1 GI:21429075
KEYWORDS FLI cDNA
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
REFERENCE Rukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 5178)
Scapleton, M., Brokstein, P., Hong, L., Agbavani, A., Carlson, J.,
Champe, M., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Frise, E.,
George, R., Gonzalez, M., Guarni, H., Kronmiller, B., Li, P., Liao, G.,
Miranda, A., Mungall, C. J., Nuncio, J., Pacle, J., Paragas, V., Park, S.,
Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M.
and Celniker, S.

Direct Submission
Submitted (10-JUN-2002) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unspliced precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our Web site
(<http://fruitfly.berkeley.edu>) or send email to
cdna@fruitfly.berkeley.edu.

FEATURES

1..5178
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="Y; cn bw sp"
/db_xref="taxon:7227"

misc_feature 1..5177
/note="sim4 alignment with AE003581 (band-23B3-band-23D3)"
466..3777
/note="Longest ORF"
/codon_start=1
/product="LD27161p"
/protein_id="AAM50257.1"
/db_xref="GI:21429076"
/translation="MNKIASNKSLIKTYNDVVKSVKVTTRIMTPASGSPVSGSGSGS
VSRASSPRASPSVSASARRSPSATPKVPPKMLAKRPSSSGKVNATPSPPHATPP
TROPLDKSGVAGSVGGGARAVKTPKSTPTTPTTTPPAKQLQSLVGRGRSY
KDSKSHSDILSLMNCASSRPSNAGSGSKILLANTQLRAAKRDVNLGISPTDL
GPPKIQTKQSTRKSSATPTPTSIHKRLNGTLASTAAIATPIKGVATKSGAT
NKLQOSAKTASTIKSPAIVGVSEESLRSGVDQTELQELANGQSATGCTPPMAEDS
KRTSINELPCETEAKIKRCEDQDQASTTASPLPMPINPQELDCARLAKFPDQTF
EPNALSILPEISNSVSTTMTLATPRVDIAAQPVLKPTFENGELDCARLAKFPDQTF
ERKEBQROIILGLSVVQFMSKELDAVACKENKHQCAKTKGTLEKILLQHTQKOCER
LPREDLEKLEWIOQOEKEYLHRTAEOKIQLQLOVETEQQHRELIKRNH
SKQAQEARMEVSHKALAKQEHRTAEOKIQLQLOVETEQQHRELIKRNH
RRSPTSLPDNASAEINLTSEADSLRCLVLELQKAEISALSKAKADLIHESBERLKL
NRVALLAEQNEMLRTELEAKTEKEIKQMEELQKAYKESIKRLLTYDKLELQTH
LKORSLQLOAESKLQDLSGSHDNSLSHSCSLRSGLEIAVTTSTSPVMKGM
FENDSVSTLEIETESFKGNTSKI VRRAGSLRNNRCPQIRQTITSHGSHGSGAT
GGSPAHNPPLSQMSATALLRSSNSGESEGRPLSRAPSKMCIKASATSAVCESS
GQRKXPQDESLADWPEDPLCCSSPQAFGIERRPRSRSTMKLMSSFAKFORIQESAG
EAMVSGANSEDESCSASSEDIMRSSASSTAGSGSLSKRPKQPPSRMSIEALPCTPM
EVWSDEADAADANGLA"

ORIGIN

Query Match 2.1%; Score 24; DB 3; Length 5178;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 800 CGCTGCTGCTGCTGCCGCGCTCGT 823
|||||
Db 216 CGCTGCTGCTGCTGCCGCGCTCGT 193

Search completed: September 15, 2004, 03:47:23
Job time : 4803.11 secs

421 TGAAGTACTTGTCTACCCCTCACACTGGACCTGGCTCCCGGATATTCTCTGGGAGCTCG 480

Qy	481	TCCTCTGGTCAGGCCATTC	CAAGCTATTATTTGTTTCGGAGCCACGACCTTCAAACTGC	541
Db	481	TCCTCTGGTCAGGCCATTC	CAAGCTATTATTTGTTTCGGAGCCACGACCTTCAAGTGC	540
Qy	541	AGTGAGCATCTCTGGCTCTG	ACCCACAGACCAATGGGACTTTGACTTTGCGTGGCTACCTG	600
Db	541	AGTGAGCATCTCTGGCTCTG	ACCCACAGACCAATGGGACTTTGACTTTGCGTGGCTACCTG	600
Qy	601	GAAGAGCCTGAAGGCCCGCGAAGTCTG	CAACTGTAAATCTCACTGTGATTGGTGTCCCA	660
Db	601	GAAGAGCCTGAAGGCCCGCGAAGTCTG	CAACTGTAAATCTCACTGTGATTGGTGTCCCA	660
Qy	661	AGACACTGGAGGTGGTATTAA	TATCCAGGTGTAATTCAAGTTTACCGAGTTTAGGTTT	720
Db	661	AGACACTGGAGGTGGTATTAA	TATCCAGGTGTAATTATCAAGTTTACCGAGTTTAGGTTT	720
Qy	721	TTCAATTCGCTACTTTGGGCAAAAGTTG	CACTTGGACTAGCAGGCACCATGCTTCTGACGCC	780
Db	721	TTCAATTCGCTACTTTGGGCAAAAGTTG	CACTTGGACTAGCAGGCACCATGCTTCTGACGCC	780
Qy	781	GACGTGTACTCTTACAATACGCTGCTG	CTGCTGCGCGCTGCTTGTGGCTGCAACTG	840
Db	781	GACGTGTACTCTTACAATACGCTGCTG	CTGCTGCGCGCTGCTTGTGTGGCTGCAACTG	840
Qy	841	CTGCTGCCGTTGTTCTGTTCTGCTGT	AGAGAAAAAGAGGATTTCGTTTCAATTTCAAAA	900
Db	841	CTGCTGCCGTTGTTCTGTTCTGCTGT	AGAGAAAAAGAGGATTTCGTTTCAATTTCAAAA	900
Qy	901	GAAATCTGAAAAGAGAGACAAA	CAAGAAACTGAGCAGAAAGTGGAAATGAAAACTC	960
Db	901	GAAATCTGAAAAGAGAGACAAA	CAAGAAACTGAGCAGAAAGTGGAAATGAAAACTC	960
Qy	961	CGGCTACAATTCAGATGAAACAAAAG	ACCACAGACACCGCTTCTCTCCCTCCCAAAATCCTG	1021
Db	961	CGGCTACAATTCAGATGAAACAAAAG	ACCACAGACACCGCTTCTCTCCCTCCCAAAATCCTG	1021

1021 TCAATCCACTGATCCTGAACAAAGAAACAGTAGCTGTGGCCCTCCTCACCGCGGCTGA 108

Db	1021	TGAATCCAGTGATCTTCAACAAGAAACAGTAGCTGTGGCCCTCTCTCACCAGCGGGCTGA	1144
Qy	1081	TCAACGTCCACCCAGGCGCAAGTCATCCACAGGCTCTTTTAATCTGGCCAGTCTCTGA	1144
Db	1081	TCAACGTCCACCCAGGCGCAAGTCATCCACAGGCTCTTTTAATCTGGCCAGTCTCTGA	1144
Qy	1141	GAAGGTCAGTAATAACAACCTAGTATAG	1168
Db	1141	GAAGGTCAGTAATAACAACCTAGTATAG	1168

RESULT 2

ABK13028	ID	ABK13028 standard; cDNA; 1175 BP.
XX	AC	ABK13028;
XX	DT	23-APR-2002 (first entry)
XX	DE	DNA encoding human B7-like protein, B7-L_hl.
XX	KW	Human; B7-like protein; B7-L; antiinfertility; gynaecological;
XX	KW	anti-tumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;
XX	KW	antiinflammatory; dermatological; antipsoriatic; neuroprotective;
XX	KW	antiadhesive; haemostatic; antithyroid; antiulcer; antiallergic;
XX	KW	antiaslathmic; nephrotropic; antibacterial; virucide; tumour; cancer;
XX	KW	reproductive disorder; graft versus host disease; autoimmune disease;
XX	KW	toxic shock syndrome; allergy; nephropathy; skin disorder;
XX	KW	endocrinopathy; lymphoproliferative disorder; gene; ss.
XX	OS	Homo sapiens.
XX	XX	

PH	Key	Location/Qualifiers	
FT	CDS	27..1175	
FT		/*tag= a	
FT		/product= "B7-like protein, B7-L_h1"	
XX			
PN	WO200200710-A2.		
PD	03-JAN-2002.		
XX			
PF	28-JUN-2001; 2001WO-US020719.		
XX			
PR	28-JUN-2000; 2000US-0214512P.		
PR	28-NOV-2000; 2000US-00729264.		
XX			
PA	(ANGE-) AMGEN INC.		
XX			
PI	Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;		
XX			
DR	WPI; 2002-130881/17.		
DR	P-PSDB; AAU75540.		
XX			
PT	New B7-like polypeptides, polynucleotides and their modulators, useful		
PT	for diagnosing, preventing and treating reproductive, immune and		
PT	proliferative disorders, e.g. cancer and arteriosclerosis.		
XX			
PS	Claim 1; Fig 1; 135pp; English.		
XX			
CC	The invention relates to an isolated B7-like (B7-L) polypeptide (I). The		
CC	polypeptide, polynucleotide encoding it and antibody against (I) are		
CC	useful for treating B7-like polypeptide-related disease, disorders or		
CC	conditions including reproductive disorders (e.g. infertility,		
CC	miscarriage, preterm labour and delivery and endometriosis) and		
CC	proliferative disorders. Antibodies, soluble proteins comprising		
CC	extracellular domains and other regulators of B7-L polypeptides are		
CC	useful for enhancing the immune response to tumours. (I) plays a role in		
CC	growth and maintenance of cancer cells based on the observation of		
CC	seminal vesicle hyperplasia in transgenic mice overexpressing B7-L		
CC	polypeptide. Hence modulators of (I) are useful for the treatment of		
CC	cancer including seminal vesicle cancer, lung, brain, breast, ovarian,		
CC	testicular cancer and cancers of haematopoietic system. B7-L polypeptide		
CC	pathway can be manipulated to regulate cytotoxic T-lymphocyte response in		
CC	allograft transplantation, graft versus host disease, T-cell dependent B-		
CC	cell mediated diseases and autoimmune diseases. B7-L molecules are useful		
CC	for alleviating the symptoms associated with diseases involving chronic		
CC	immune cell dysfunction or to treat autoimmune diseases such as systemic		
CC	lupus erythematosus, rheumatoid arthritis, multiple sclerosis, diabetes,		
CC	immune thrombocytopenic purpura and psoriasis, chronic inflammatory		
CC	disease such as inflammatory bowel disease (Crohn's disease and		
CC	ulcerative colitis), Grave's disease, Hashimoto's thyroiditis and		
CC	diabetes mellitus. They are also useful as immunosuppressive agents for		
CC	bone marrow and organ transplantation or to prolong graft survival. B7-L		
CC	molecules are also useful for diagnosis and treatment of diseases		
CC	involving abnormal cell proliferation, including arteriosclerosis and		
CC	vascular stenosis. Antagonists of B7-L polypeptides are useful for		
CC	alleviation of toxic shock syndrome or allosensitisation due to blood		
CC	transfusions, and for treatment of allergy, asthma and hypersensitivity		
CC	reactions, nephropathies (e.g. glomerulonephritis), skin disorders		
CC	(pemphigus and pemphigoid), endocrinopathies (Grave's disease), various		
CC	pneumopathies (extrinsic alveolitis), vasculopathies, coeliac disease,		
CC	anaemia, thrombocytopenias, Guillain-Barre syndrome and myasthenia		
CC	gravis, and lymphoproliferative disorders such as multiple myeloma. The		
CC	present sequence represents the coding sequence of human B7-L_h1		
XX			
SQ	Sequence 1175 BP; 295 A; 307 C; 286 G; 287 T; 0 U; 0 Other;		
	Query Match	87.9%; Score 1027; DB 6; Length 1175;	
	Best Local Similarity	99.9%; Pred. No. 0;	
	Matches 1077; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
QY	91	ACTCTCTGAAGGGCTCCAGGCTCGCTTCAACATGACCGCTCCAGGGCTGAAGCTCAT	150
DB	98	AGTCTCTGAAGGGCTCCAGGCTCGCTTCAACATGACCGCTCCAGGGCTGAAGCTCAT	157

QY	151	CATGTGGGCTCTCAGTGACATGGTGGTGGTAAAGCGTCAGGCCCATGGAGCCCATCATCAC	210
DB	158	CATGTGGGCTCTCAGTGACATGGTGGTAAAGCGTCAGGCCCATGGAGCCCATCATCAC	217
QY	211	CAATGACCGCTTCACTCTCAGAGGTACGACACAGGGCGGGAATCTTCACTCCGAGATGAT	270
DB	218	CAATGACCGCTTCACTCTCAGAGGTACGACACAGGGCGGGAATCTTCACTCCGAGATGAT	277
QY	271	CATCCACAATGTGGAGCCAGTGTATGGGGAACATCAGATGACGCTCCAGAACAGTCG	330
DB	278	CATCCACAATGTGGAGCCAGTGTATGGGGAACATCAGATGACGCTCCAGAACAGTCG	337
QY	331	CCTGATGATCTGCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTCAATCCAGTGT	390
DB	338	CCTGATGATCTGCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTCAATCCAGTGT	397
QY	391	TAATCTTTGTAGTCGCTGAGAATGAACCTTGTGAAGTTACTTGTCTACCTCAGCTGAC	450
DB	398	TAATCTTTGTAGTCGCTGAGAATGAACCTTGTGAAGTTACTTGTCTACCTCAGCTGAC	457
QY	451	CTGCTCCCGGATATTTCTGGGAGTCGGTCTCTCTGTCAGCCATTCAGCTATTATTT	510
DB	458	CCGGCTCCCGGATATTTCTGGGAGTCGGTCTCTCTGTCAGCCATTCAGCTATTATTT	517
QY	511	TGTTCCCGAGCCGAGCGACCTTCAAGTCGAGTCAGCATCTGCTCTGACCCACAGAG	570
DB	518	TGTTCCCGAGCCGAGCGACCTTCAAGTCGAGTCAGCATCTGCTCTGACCCACAGAG	577
QY	571	CAATGGACTTTGACTTGGCTGGCTACTCTGGAAGAGCTGGAAGCCCGCAAGTCTGCAAC	630
DB	578	CAATGGACTTTGACTTGGCTGGCTACTCTGGAAGAGCTGGAAGCCCGCAAGTCTGCAAC	637
QY	631	TGTAATCTCAGTCGTGATTCGGTGTCCCAAGACACTGGAGGTGGTATTATATTTCCAGG	690
DB	638	TGTAATCTCAGTCGTGATTCGGTGTCCCAAGACACTGGAGGTGGTATTATATTTCCAGG	697
QY	691	TGTATTATCAAGTTTACCGAGTTTAGTGTTCCTTTCATTCGCTTACTTGGGGCAAGTTGACT	750
DB	698	TGTATTATCAAGTTTACCGAGTTTAGTGTTCCTTTCATTCGCTTACTTGGGGCAAGTTGACT	757
QY	751	TGAGCTAGCAGGACCATGCTTCTGACGGCGAGCTGTACTTCTTACAATACGCTGCTGCTG	810
DB	758	TGAGCTAGCAGGACCATGCTTCTGACGGCGAGCTGTACTTCTTACAATACGCTGCTGCTG	817
QY	811	CTGCCCGCTGCTGTGTGTGGCTGMAACTGCTGCTGCCGTTGTTTCTGCTGTAGAAG	870
DB	818	CTGCCCGCTGCTGTGTGTGGCTGMAACTGCTGCTGCCGTTGTTTCTGCTGTAGAAG	877
QY	871	AAAAAGAGGATTTCTGATTTCAATTTCAAAAGAAATCTGAAAGAGAGACAACAAGA	930
DB	878	AAAAAGAGGATTTCTGATTTCAATTTCAAAAGAAATCTGAAAGAGAGACAACAAGA	937
QY	931	AAGTACAGAGAAAGTGGAAATGAAATCTCCGCTACAATTCAGATGAACAAAGACAC	990
DB	938	AAGTACAGAGAAAGTGGAAATGAAATCTCCGCTACAATTCAGATGAACAAAGACAC	997
QY	991	AGACACCGCTTCTCTCCCTCCCAATCTGTGAATCCAGTGTCTTGAACAAAGAACAG	1050
DB	998	AGACACCGCTTCTCTCCCTCCCAATCTGTGAATCCAGTGTCTTGAACAAAGAACAG	1057
QY	1051	TAGCTGTGGCCCTCTCCACGAGGGCTGATCAAGTCCACCCAGGCGCAGAGTCATCC	1110
DB	1058	TAGCTGTGGCCCTCTCCACGAGGGCTGATCAAGTCCACCCAGGCGCAGAGTCATCC	1117
QY	1111	ACAGGCTCTTTTAAATCTGGCCAGTCTCTGAGAAGTCTAGTATACAACTGTAGTATAG	1168
DB	1118	ACAGGCTCTTTTAAATCTGGCCAGTCTCTGAGAAGTCTAGTATACAACTGTAGTATAG	1175

RESULT 3
ABK13030
ID ABK13030 standard; cDNA; 1240 BP.
XX

QY 991 AGACACCGCTTCTCTCCCTCCCAATCTCTGTGAATCCAGTGATCTCTGAAACAAGAAACAG 1050
 Db |||||
 1063 AGAAACCGCTTCTCTCCCTCCCAATCTCTGTGAATCCAGTGATCTCTGAAACAAGAAACAG 1122
 QY 1051 TAGCTGTGGCCCTCTCTACACAGCGGCTGATCAACAGTCCACCCAGGCCAGCAAGTCATCC 1110
 Db |||||
 1123 TAGCTGTGGCCCTCTCTACACAGCGGCTGATCAACAGTCCACCCAGGCCAGCAAGTCATCC 1182
 QY 1111 ACAGGCTCTTTTAATCTGCGCAGTCTCTGAGAGGTCAGTAAATACAACCTGTAGTATAG 1168
 Db |||||
 1183 ACAGGCTCTTTTAATCTGCGCAGTCTCTGAGAGGTCAGTAAATACAACCTGTAGTATAG 1240

RESULT 4

ADB62950

ID ADB62950 standard; cDNA; 2051 BP.

AC ADB62950;

XX 04-DEC-2003 (first entry)

XX Human cDNA encoding clone PLACE60177880.

XX Human; ss; Gene; pharmaceutical; diagnostic; gene therapy;
 KW tissue regeneration; cell regeneration; membrane protein;
 KW signal transduction-related protein; transcription-related protein;
 KW osteoporosis; neurological disease; cancer; tumour.

XX Homo sapiens.

XX Key Location/Qualifiers

FH 104..1327

FT /*tag= a

FT /product= "Clone PLACE60177880 protein"

XX EP1308459-A2.

XX 07-MAY-2003.

XX 28-MAR-2002; 2002EP-00007401.

XX 05-NOV-2001; 2001JP-00379298.

XX 25-JAN-2002; 2002US-00350978.

XX (HELI-) HELIX RES INST.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuho Y;

XX WPI; 2003-450961/43.

XX P-PSDB; ADB64920.

XX New polynucleotides and polypeptides, useful for developing a diagnostic
 PT marker or medicines for regulation of their expression and activity, or
 PT as targets of gene therapy.

XX Claim 1; Page; 222pp; English.

XX The invention discloses a polynucleotide comprising a sequence selected
 CC from 1970 fully defined nucleotide sequences which encode novel
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
 CC or its partial peptide, an antibody binding to the polypeptide or peptide
 CC of the polynucleotide, immunologically assaying the polypeptide or
 CC peptide of the polynucleotide by contacting the polypeptide or
 CC with the antibody of the encoded protein, and observing the binding
 CC between the two, a transformant carrying the polynucleotide in an
 CC expressible manner and an antisense polynucleotide. The oligonucleotide
 CC is useful as a primer for synthesising the polynucleotide, or as a probe
 CC for detecting the polynucleotide. The polynucleotides and encoded
 CC proteins are useful as pharmaceutical agents and many disease-related
 CC genes may be included in them, for developing a diagnostic marker or

CC medicines for regulation of their expression and activity, or as targets
 CC of gene therapy. The genes are involved in tissue and/or cell
 CC regeneration. Membrane proteins, signal transduction-related genes,
 CC transcription-related proteins, disease-related proteins and genes
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
 CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
 CC the activity or expression of the encoded protein to treat diseases. The
 CC sequence presented is a cDNA of the invention. Note: Some of the sequence
 CC data for this patent is not represented in the printed specification, but
 CC is based on sequence information supplied by the European Patent Office.

XX

SQ Sequence 2051 BP; 580 A; 463 C; 474 G; 534 T; 0 U; 0 Other;

Query Match

Best Local Similarity 83.6%; Score 976; DB 9; Length 2051;

Matches 1076; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 91 AGTCTGAAGGGCTCCACAGGCTCGCTTCAACTGACCGTCTCCAGGGCTGGAAGCTCAT 150

Db |||||

250 AGTCTGAAGGGCTCCACAGGCTCGCTTCAACTGACCGTCTCCAGGGCTGGAAGCTCAT 309

QY 151 CATGTGGGCTCTCAGTGACATGGTGGTCTAAGCGTCAGGCCCATGGAGCCCATCATCAC 210

Db |||||

310 CATGTGGGCTCTCAGTGACATGGTGGTCTAAGCGTCAGGCCCATGGAGCCCATCATCAC 369

QY 211 CAATGACCGCTTCACTCTCAGAGGTACGACAGGGGGGAATTCACCTCGGAGATGAT 270

Db |||||

370 CAATGACCGCTTCACTCTCAGAGGTACGACAGGGGGGAATTCACCTCGGAGATGAT 429

QY 271 CATCCACAATGTGGAGCCAGTGAATCGGGGAACATCAGATGCGAGCCTCCAGAACAGTCG 330

Db |||||

430 CATCCACAATGTGGAGCCAGTGAATCGGGGAACATCAGATGCGAGCCTCCAGAACAGTCG 489

QY 331 CTGTCATGATTCGCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCAATTCACAGTGT 390

Db |||||

490 CTGTCATGATTCGCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCAATTCACAGTGT 549

QY 391 TAATCTTGTAGTCGCTGAGAAATGAACCTTGTGAAGTTACTTGTCTACCCCTCACACTGGAC 450

Db |||||

550 TAATCTTGTAGTCGCTGAGAAATGAACCTTGTGAAGTTACTTGTCTACCCCTCACACTGGAC 609

QY 451 CTGGCTCCCGGATATTTCTGGGAGCTCGCTCTCCCTGCTCAGCCATTCAGCTATTATTT 510

Db |||||

610 CCGGCTCCCGGATATTTCTGGGAGCTCGCTCTCCCTGCTCAGCCATTCAGCTATTATTT 669

QY 511 TGTTCGGAGCCAGCGACCTTCAAAGTGAGTGAGCATCTCGCTCTGACCCCAACAGAG 570

Db |||||

670 TGTTCGGAGCCAGCGACCTTCAAAGTGAGTGAGCATCTCGCTCTGACCCCAACAGAG 729

QY 571 CAATGGAGCTTTGACCTTGGTGGCTACCTGGAAGAGCCTGAAAGCCCGCAAGTCTGCAAC 630

Db |||||

730 CAATGGAGCTTTGACCTTGGTGGCTACCTGGAAGAGCCTGAAAGCCCGCAAGTCTGCAAC 789

QY 631 TGTAAATCTCACTGTGATTCGGTGTCCCAAGACACTGGAGGTGTTAATAATTCCAG 690

Db |||||

790 TGTAAATCTCACTGTGATTCGGTGTCCCAAGACACTGGAGGTGTTAATAATTCCAG 849

QY 691 TGTATTATCAAGTTTACCGAGTTTAGGTTTTCATTGCTTCTGGGCAAGTCTGGACT 750

Db |||||

850 TGTATTATCAAGTTTACCGAGTTTAGGTTTTCATTGCTTCTGGGCAAGTCTGGACT 909

QY 751 TGGACTAGCAGCAGCATGCTTCTGACGCCAGCTGTACTTTACAATACGCTGCTGCTG 810

Db |||||

910 TGGACTAGCAGCAGCATGCTTCTGACGCCAGCTGTACTTTACAATACGCTGCTGCTG 969

QY 811 CTGCGCGCTGCTGTTGTGGCTGCAACTGCTGCGCTTGTGTTCTGCTGTAGAG 870

Db |||||

970 CTGCGCGCTGCTGTTGTGGCTGCAACTGCTGCGCTTGTGTTCTGCTGTAGAG 1029

QY 871 AAAAAAGAGATTCGTATTTCATTTCAAAAGAAATCTGAAAAGAGAGACAAAAGAG 930

Db |||||

1030 AAAAAAGAGATTCGTATTTCATTTCAAAAGAAATCTGAAAAGAGAGACAAAAGAG 1089

Wed Sep 15 10:41:07 2004

```

13-OCT-2003 (first entry)
Human adult heart cDNA #444.
Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
genome mapping; biodiversity; genetic disorder.
Homo sapiens.
US2003073623-A1.
17-APR-2003.
30-JUL-2001; 2001US-00918995.
30-JUL-2001; 2001US-00918995.
(DRNA/) DRMANAC R T.
(LABA/) LABAT I.
(STAC/) STACHE-CRAIN B.
(DICK/) DICKSON M C.
(JONE/) JONES L W.
Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
WPI; 2003-615964/58.
New polynucleotide sequences obtained from various cDNA libraries, useful
as hybridization probes, as oligomers for PCR, for chromosome and gene
mapping, in the recombinant production of protein, or in generating
antisense DNA or RNA.
Claim 1; SEQ ID NO 3342; 44pp; English.
The invention relates to an isolated polynucleotide comprising any one of
38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
determined by the technique of SBH (sequencing by hybridisation). Also
included is a purified polypeptide comprising a sequence corresponding to
a reading frame of the novel polynucleotide. The nucleic acid sequences
are useful in diagnostics as expressed sequence tags (EST) for
identifying expressed genes or for physical mapping of the human genome,
in forensics, in assessing biodiversity, or in identifying mutations
responsible for genetic disorders and other traits. The nucleotide
sequences are also useful as hybridisation probes, as oligomers for PCR,
for chromosome and gene mapping, in the recombinant production of
protein, or in generating antisense DNA or RNA. The purified polypeptide
is useful for generating antibodies specific for it. The present sequence
is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
for this patent did not form part of the printed specification, but was
obtained in electronic format directly from USPTO at
seqdata.uspto.gov/sequence.html?docid=20030073623
Sequence 474 BP; 154 A; 128 C; 96 G; 95 T; 0 U; 1 Other;
Query Match 18.6%; Score 217; DB 8; Length 474;
Best Local Similarity 19.6%; Pred. No. 3e-99;
Matches 267; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 901 GAAATCTGAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGAATGAAATCTC 960
DB 63 GAAATCTGAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGAATGAAATCTC 122
QY 961 GGGTACATTCAGATGACAAAGAACACACAGACACCGCTCTCTCCCTCCCAATCTG 1020
DB 123 GGGTACATTCAGATGACAAAGAACACACAGACACCGCTCTCTCCCTCCCAATCTG 182
QY 1021 TGAATCCAGTATCTGACAAAGAAACAGTAGCTGTGSCCTCTCTCCACAGCGGCTGA 1080
DB 183 TGAATCCAGTATCTGACAAAGAAACAGTAGCTGTGSCCTCTCTCCACAGCGGCTGA 242
QY 1081 TCAAGCTCCACCCAGGACAGCAAGTCATCCAGAGGCTTTTAAATCTGGCCAGCTCGA 1140
DB 243 TCAAGCTCCACCCAGGACAGCAAGTCATCCAGAGGCTTTTAAATCTGGCCAGCTCGA 302
1141 GAAGTCAGTAATACTACTGTAGTATAG 1168
303 GAAGTCAGTAATACTACTGTAGTATAG 330
RESULT 8
AAI36582
ID AAI36582 standard; DNA; 401 BP.
XX
AC AAI36582;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #5268 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488897/53.
XX
DR Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 5268; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENPs).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 401 BP; 105 A; 98 C; 111 G; 87 T; 0 U; 0 Other;
Query Match 16.4%; Score 191; DB 4; Length 401;
Best Local Similarity 100.0%; Pred. No. 3.4e-85;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 44 GGTTCGTGGTCTGGTAAATGAAAGTCATAGAGGCCCCCAAAATGCAAGAGTCTTGAGGGC 103
DB 211 GGTTCGTGGTCTGGTAAATGAAAGTCATAGAGGCCCCCAAAATGCAAGAGTCTTGAGGGC 270
QY 104 TCCAGAGGTCGGTTCACCTGACCGTCTCCAGGGCTGGAAGTCATCATGTGGGCTCTC 163
DB 271 TCCAGAGGTCGGTTCACCTGACCGTCTCCAGGGCTGGAAGTCATCATGTGGGCTCTC 330
QY 164 AGTGACATGGTGTCTTAAGGTCAGGCCCATGGAGCCCATCATCAACCAATGACCGCTTC 223
DB 331 AGTGACATGGTGTCTTAAGGTCAGGCCCATGGAGCCCATCATCAACCAATGACCGCTTC 390
224 ACCTCTCAGAG 234
391 ACCTCTCAGAG 401
```


RESULT_15	
AAV35364	
ID	AAV35364 standard; cDNA; 1200 BP.
XX	
XX	
AAV35364;	
XX	
28-SEP-1998	(first entry)
XX	
Human GDNF alpha-3 receptor cDNA #1.	
XX	
Glial cell line-derived neurotrophic factor alpha-3 receptor; GDNF;	
KW	treatment; neurodegenerative disease; Parkinson's Disease; ALS; SMA;
KW	amyotrophic lateral sclerosis; spinal muscular atrophy; nerve; trauma;
KW	Huntington's Disease; Alzheimer's Disease; diabetic neuropathy; muscle;
KW	muscular dystrophy; diagnostic; ss.
XX	
Homo sapiens.	
XX	
XX	
Key	Location/Qualifiers
CDS	1..1200
FT	/*tag= a
FT	/product= "GDNF alpha-3"
FT	/note= "partial sequence of glial cell-derived
FT	neurotrophic factor alpha-3 receptor"
XX	
EP846764-A2.	
XX	
10-JUN-1998.	
XX	
20-NOV-1997;	97EP-00309375.
XX	
27-NOV-1996;	96GB-00024677.
PR	
09-MAY-1997;	97GB-00009463.
XX	
(SMK) SMITHKLINE BEECHAM PLC.	
XX	
Lawrence GMP;	
XX	
WPI; 1998-299980/27.	
DR	
P-PSDB; AAW65116.	
XX	
New factor alpha 3 receptor polypeptide and e.g. DNA and agonists - used	
PT	to treat neuro degenerative diseases, muscular diseases and nerve and
PT	muscle trauma and in diagnostic assays.
XX	
Claim 9; Fig 1; 22pp; English.	
XX	
This sequence encodes a novel glial cell line-derived neurotrophic factor	
CC	alpha-3 receptor (GDNF alpha-3). This protein can be used to treat e.g.
CC	neurodegenerative diseases (such as Parkinson's Disease, amyotrophic
CC	lateral sclerosis (ALS), spinal muscular atrophy (SMA), Huntington's
CC	Disease, Alzheimer's Disease, diabetic neuropathy), muscular diseases
CC	(including the muscular dystrophies) and nerve and muscle trauma and in
CC	diagnostic assays for such conditions
XX	
Sequence 1200 BP; 236 A; 419 C; 307 G; 238 T; 0 U; 0 Other.	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 23:14:24 ; Search time 91.2755 Seconds
(without alignments)
7101.389 Million cell updates/sec

Title: US-09-729-264-3

Perfect score: 1168

Sequence: 1 agtgatcatgtggcaggag.....gtaatacaactgtagtatag 1168

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1281558

Minimum DB seq length: 16

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	1.9	1203	3	US-09-220-528-65
2	22	1.9	1699	4	US-09-187-906-20
3	22	1.9	3942	3	US-09-162-484-19
4	21	1.8	550	4	US-09-669-751-128
5	21	1.8	1428	4	US-09-489-039A-3243
6	20	1.7	394	4	US-09-621-976-16752
7	20	1.7	794	4	US-09-621-976-16751
8	20	1.7	984	4	US-09-252-991A-12897
9	20	1.7	1857	4	US-09-252-991A-12897
10	20	1.7	1910	3	US-09-593-711A-3
11	20	1.7	1914	1	US-07-601-094-1
12	20	1.7	1914	1	US-08-012-735-1
13	20	1.7	2214	3	US-08-864-038A-1
14	20	1.7	3331	3	US-08-864-038A-2
15	20	1.7	3331	3	US-08-864-038A-4
16	20	1.7	6407	2	US-08-616-844-7
17	20	1.7	6407	2	US-08-599-654-7
18	20	1.7	6407	3	US-08-944-868A-7
19	20	1.7	6407	3	US-08-944-423A-7
20	20	1.7	6407	3	US-08-944-496-7
21	19	1.6	402	4	US-09-621-976-88
22	19	1.6	1425	1	US-08-464-148-1
23	19	1.6	1425	1	US-08-385-500-1
24	19	1.6	1425	1	US-08-846-784-1
25	19	1.6	1477	4	US-09-620-312D-1019
26	19	1.6	1768	4	US-09-833-381-523
27	19	1.6	1899	4	US-09-919-060-15

```

c 28 19 1.6 1899 4 US-09-919-060-16 Sequence 16, Appli
29 19 1.6 1899 1 US-08-447-500-5 Sequence 5, Appli
30 19 1.6 1899 1 US-08-454-097-5 Sequence 5, Appli
31 19 1.6 1899 1 US-08-453-866-5 Sequence 5, Appli
32 19 1.6 1899 3 US-08-185-359-5 Sequence 5, Appli
33 19 1.6 1995 4 US-09-620-312D-908 Sequence 908, App
34 19 1.6 2023 2 US-09-491-522-6 Sequence 6, Appli
35 19 1.6 2115 2 US-08-474-379C-60 Sequence 60, Appl
36 19 1.6 2115 3 US-09-146-249A-60 Sequence 60, Appl
37 19 1.6 2115 3 US-08-206-188B-60 Sequence 60, Appl
38 19 1.6 2439 4 US-09-489-039A-4904 Sequence 4904, Ap
39 19 1.6 2450 4 US-09-491-522-2 Sequence 2, Appli
40 19 1.6 2617 4 US-09-786-240-21 Sequence 21, Appli
41 19 1.6 2693 4 US-09-919-060-12 Sequence 12, Appli
42 19 1.6 2693 4 US-09-919-060-14 Sequence 14, Appli
43 19 1.6 3311 4 US-09-367-891A-5 Sequence 5, Appli
44 19 1.6 5092 3 US-09-412-545-1 Sequence 1, Appli
45 19 1.6 5621 4 US-09-566-921-106 Sequence 106, App

```

ALIGNMENTS

```

RESULT 1
US-09-220-528-65
; Sequence 65, Application US/09220528A
; Patent No. 6284540
; GENERAL INFORMATION:
; APPLICANT: Milbrandt, Jeffrey D.
; TITLE OF INVENTION: Artemin, A No. 6284540el Neurotrophic Factor
; FILE REFERENCE: 6029-7998
; CURRENT APPLICATION NUMBER: US/09/220,528A
; EARLIER FILING DATE: 1998-12-24
; EARLIER FILING DATE: 1998-12-24
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 60/108,148
; EARLIER FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/163,283
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 65
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-528-65

Query Match 1.9% Score 22; DB 3; Length 1203;
Best Local Similarity 100.0%; Pred.No. 0.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCTGCTGCCCGCTCG 822
|||||
Db 51 GCTGCTGCTGCTGCTGCCCGCTCG 72

RESULT 2
US-09-187-906-20
; Sequence 20, Application US/09187906
; Patent No. 667135
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: Ret Ligand (RetL) for Stimulating Neural
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02142

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/187,906
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/07726
FILING DATE: 07-MAY-97
APPLICATION NUMBER: US 60/017,427
FILING DATE: 08-MAY-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/019,300
FILING DATE: 07-JUN-96
APPLICATION NUMBER: US 60/021,859
FILING DATE: 16-JUL-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/043,533
FILING DATE: 10-APR-97
ATTORNEY/AGENT INFORMATION:
NAME: Kaplan, Warren A.
REGISTRATION NUMBER: 34,199
REFERENCE/DOCKET NUMBER: A008 PCT CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-679-2400
TELEFAX: 617-679-2838
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 1699 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 175..1374
US-09-187-906-20

Query Match 1.9%; Score 22; DB 4; Length 1699;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCTGCCGCGTCG 822
Db 225 GCTGCTGCTGCTGCCGCGTCG 246

RESULT 3
US-09-162-484-19
Sequence 19, Application US/09162484
Patent No. 6248724
GENERAL INFORMATION:
APPLICANT: Phillips, M. Ian
TITLE OF INVENTION: MOHUCZY, Dagmara
TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME RNA AND METHODS OF USE
FILE REFERENCE: UFLA/087/UFLA087P
CURRENT APPLICATION NUMBER: US/09/162,484
CURRENT FILING DATE: 1998-09-25
EARLIER APPLICATION NUMBER: 60/059,661
EARLIER FILING DATE: 1997-09-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 3942
TYPE: DNA
ORGANISM: Rattus norvegicus
US-09-162-484-19

Query Match 1.9%; Score 22; DB 3; Length 3942;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCTGCCGCGTCG 822
Db 72 GCTGCTGCTGCTGCCGCGTCG 93

RESULT 4
US-09-669-751-128/C
Sequence 128, Application US/09669751
Patent No. 6551575
GENERAL INFORMATION:
APPLICANT: Greenspan, Ralph J.
TITLE OF INVENTION: Methods for Identifying Compounds for
TITLE OF INVENTION: Motion Sickness, Vertigo and Other Disorders Related to
TITLE OF INVENTION: Balance and the Perception of Gravity
FILE REFERENCE: P-NI 3864
CURRENT APPLICATION NUMBER: US/09/669,751
CURRENT FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US 60/168,579
PRIOR FILING DATE: 1999-12-02
NUMBER OF SEQ ID NOS: 261
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 128
LENGTH: 550
TYPE: DNA
ORGANISM: Drosophila
US-09-669-751-128

Query Match 1.8%; Score 21; DB 4; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.64;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 837 ACTGCTGCTGCCGTTGTTGTT 857
Db 317 ACTGCTGCTGCCGTTGTTGTT 297

RESULT 5
US-09-489-039A-3243/C
Sequence 3243, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Bregon et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 3243
LENGTH: 1428
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3243

Query Match 1.8%; Score 21; DB 4; Length 1428;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 GCTTCAACTGCACCGTCTCCC 134
Db 1147 GCTTCAACTGCACCGTCTCCC 1127

RESULT 6
US-09-621-976-16752
Sequence 16752, Application US/09621976
Patent No. 6639063

```
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16752
; LENGTH: 394
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16752

Query Match      1.7%; Score 20; DB 4; Length 394;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      810 GCTGCCGCCGCTGTTGTGT 829
Db      34 GCTGCCGCCGCTGTTGTGT 53

RESULT 7
US-09-621-976-16751
; Sequence 16751, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16751
; LENGTH: 794
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16751

Query Match      1.7%; Score 20; DB 4; Length 794;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      810 GCTGCCGCCGCTGTTGTGT 829
Db      34 GCTGCCGCCGCTGTTGTGT 53

RESULT 8
US-09-252-991A-12897
; Sequence 12897, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; LENGTH: 984
; TYPE: DNA
US-09-252-991A-12897

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12897

Query Match      1.7%; Score 20; DB 4; Length 984;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      827 TGTGGCTGCAACTGCTGCTG 846
Db      395 TGTGGCTGCAACTGCTGCTG 414

RESULT 9
US-09-252-991A-12735
; Sequence 12735, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12735
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12735

Query Match      1.7%; Score 20; DB 4; Length 1857;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      827 TGTGGCTGCAACTGCTGCTG 846
Db      1746 TGTGGCTGCAACTGCTGCTG 1765

RESULT 10
US-09-593-711A-3/c
; Sequence 3, Application US/09593711A
; Patent No. 6271030
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Madeline M. Butler
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF C/EBP BETA EXPRESSION
; FILE REFERENCE: RTS-0118
; CURRENT APPLICATION NUMBER: US/09/593,711A
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 244
; SEQ ID NO 3
; LENGTH: 1910
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
; NAME/KEY: unsure
; LOCATION: 1415
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1421
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1422
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1423
```

Wed Sep 15 10:41:07 2004

OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 1424
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 1458
OTHER INFORMATION: unknown
NAME/KEY: CDS
LOCATION: (299)...(1336)
US-09-593-711A-3

Query Match 1.7%; Score 20; DB 3; Length 1910;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 800 CGCTGCTGCTGCTGCCGCG 819
Db 205 CGCTGCTGCTGCTGCCGCG 186

RESULT 11
US-07-601-094-1/c
Sequence 1, Application US/07601094
Patent No. 5215892
GENERAL INFORMATION:
APPLICANT: Kishimoto, Tadimitsu
APPLICANT: Hirano, Toshio
APPLICANT: Akira, Shizuo
APPLICANT: Isshiki, Hiroshi
APPLICANT: Tanabe, Osamu
APPLICANT: Kinoshita, Shigemi
APPLICANT: Shimamoto, Takuva
TITLE OF INVENTION: C/EBP2 Gene and Recombinant
TITLE OF INVENTION: C/EBP2
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak &
ADDRESSEE: Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/601,094
FILING DATE: 19901022
CLASSIFICATION: 435
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1914 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 281..1316
OTHER INFORMATION:

Query Match 1.7%; Score 20; DB 1; Length 1914;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 800 CGCTGCTGCTGCTGCCGCG 819
Db 204 CGCTGCTGCTGCTGCCGCG 185

RESULT 13
US-08-864-038A-1
Sequence 1, Application US/08864038A
Patent No. 6001592
GENERAL INFORMATION:
APPLICANT: Kunio NAKASHIMA et al.
TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR

Query Match 1.7%; Score 20; DB 1; Length 1914;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 800 CGCTGCTGCTGCTGCCGCG 819
Db 204 CGCTGCTGCTGCTGCCGCG 185

RESULT 12
US-08-012-735-1/c
Sequence 1, Application US/08012735
Patent No. 5360894
GENERAL INFORMATION:
APPLICANT: Kishimoto, Tadimitsu
APPLICANT: Hirano, Toshio
APPLICANT: Akira, Shizuo
APPLICANT: Isshiki, Hiroshi
APPLICANT: Tanabe, Osamu
APPLICANT: Kinoshita, Shigemi
APPLICANT: Shimamoto, Takuva
TITLE OF INVENTION: C/EBP2 Gene and Recombinant
TITLE OF INVENTION: C/EBP2
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak &
ADDRESSEE: Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/012,735
FILING DATE: 19930203
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/601,094
FILING DATE: 22 OCT 1990
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1914 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 281..1316
OTHER INFORMATION:

Query Match 1.7%; Score 20; DB 1; Length 1914;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 800 CGCTGCTGCTGCTGCCGCG 819
Db 204 CGCTGCTGCTGCTGCCGCG 185

RESULT 13
US-08-864-038A-1
Sequence 1, Application US/08864038A
Patent No. 6001592
GENERAL INFORMATION:
APPLICANT: Kunio NAKASHIMA et al.
TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR

```
; TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
; TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
; TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
; TITLE OF INVENTION: TO SAID POLYPEPTIDE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 812-5 Hirano
; STREET: Isshinden
; CITY: Tsu-city
; STATE: Mie-prefecture
; COUNTRY: JAPAN
; ZIP: 514-01
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Microsoft Windows 95
; SOFTWARE: Word Perfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,038A
; FILING DATE: May 28, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-184459
; FILING DATE: 15-July-1996
; REFERENCE/DOCKET NUMBER: F-5610
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)986-2340
; TELEFAX: (212)953-7733
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2214
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; ORGANISM: Pinctada fucata
; CELL TYPE: mantle epithelial cell
; US-08-864-038A-1

Query Match 1.7%; Score 20; DB 3; Length 2214;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 800 CGCTGCTGCTGCTGCCGCCG 819
Db 855 CGCTGCTGCTGCTGCCGCCG 874

RESULT 14
US-08-864-038A-2
; Sequence 2, Application US/08864038A
; Patent No. 6001592
; GENERAL INFORMATION:
; APPLICANT: Kunio NAKASHIMA et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
; TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
; TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
; TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 812-5 Hirano
; STREET: Isshinden
; CITY: Tsu-city
; STATE: Mie-prefecture
; COUNTRY: JAPAN
; ZIP: 514-01
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Microsoft Windows 95
```

```
; SOFTWARE: Word Perfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,038A
; FILING DATE: May 28, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-184459
; FILING DATE: 15-July-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-5610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)986-2340
; TELEFAX: (212)953-7733
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3331
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Pinctada fucata
; CELL TYPE: mantle epithelial cell
; FEATURE: mRNA
; LOCATION: from 1 to 3331
; IDENTIFICATION METHOD: E (by experiment)
; US-08-864-038A-2

Query Match 1.7%; Score 20; DB 3; Length 3331;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 800 CGTGTGCTGCTGCCGCCG 819
Db 904 CGTGTGCTGCTGCCGCCG 923

RESULT 15
US-08-864-038A-4
; Sequence 4, Application US/08864038A
; Patent No. 6001592
; GENERAL INFORMATION:
; APPLICANT: Kunio NAKASHIMA et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
; TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
; TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
; TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 812-5 Hirano
; STREET: Isshinden
; CITY: Tsu-city
; STATE: Mie-prefecture
; COUNTRY: JAPAN
; ZIP: 514-01
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Microsoft Windows 95
; SOFTWARE: Word Perfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,038A
; FILING DATE: May 28, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-184459
; FILING DATE: 15-July-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-5610
; TELECOMMUNICATION INFORMATION:
```

Wed Sep 15 10:41:07 2004

TELEPHONE: (212)986-2340
TELEFAX: (212)953-7733
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3331
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Pinictada fucata
CELL TYPE: mantle epithelial cell
FEATURE:
NAME/KEY: CDS
LOCATION: from 50 to 2263
IDENTIFICATION METHOD: P (by similarity to some other pattern)
US-08-864-038A-4
Query Match 1.7%; Score 20; DB 3; Length 3331;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 800 CGCTGCTGCTGCTGCCGCCG 819
|||||
Db 904 CGCTGCTGCTGCTGCCGCCG 923
Search completed: September 15, 2004, 06:48:27
Job time : 92.2755 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 23:16:34 ; Search time 620.673 Seconds
(without alignments)
9468.459 Million cell updates/sec

Title: US-09-729-264-3

Perfect score: 1168

Sequence: 1 agtgatcatgggtggcaggag.....gtaatacaactgtatag 1168

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 3304383 seqs, 2515761380 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6509420

Minimum DB seq length: 16

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10C_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	976	83.6	2051	16 US-10-104-047-1104	Sequence 1104, Ap
2	217	18.6	474	10 US-09-918-995-3342	Sequence 3342, Ap
3	191	16.4	401	9 US-09-864-761-16305	Sequence 16305, A
4	25	2.1	357	9 US-09-864-761-16653	Sequence 16653, A
5	23	2.0	1284	17 US-10-437-963-81852	Sequence 81852, A
6	22	1.9	767	9 US-09-770-445-930	Sequence 930, Ap
7	22	1.9	1203	9 US-09-220-920-65	Sequence 65, Appl
8	22	1.9	1829	9 US-09-828-366-15	Sequence 15, Appl
9	22	1.9	4142	12 US-10-152-319A-1498	Sequence 1498, Ap
10	21	1.8	462	17 US-10-437-963-41506	Sequence 41506, A
11	21	1.8	493	10 US-09-918-995-11051	Sequence 11051, A
12	21	1.8	550	15 US-10-255-536-128	Sequence 128, Ap
13	21	1.8	602	17 US-10-767-701-28842	Sequence 28842, A
14	21	1.8	1558	10 US-09-930-213-20	Sequence 20, Appl

C 15	21	1.8	3362	13 US-10-087-192-185	Sequence 185, App
C 16	21	1.8	94917	13 US-10-087-192-184	Sequence 184, App
C 17	20	1.7	330	17 US-10-437-963-55684	Sequence 55684, A
C 18	20	1.7	358	13 US-09-823-245A-58	Sequence 58, Appl
C 19	20	1.7	398	9 US-09-983-965-4945	Sequence 4945, Ap
C 20	20	1.7	412	10 US-09-918-995-5855	Sequence 5855, Ap
C 21	20	1.7	507	17 US-10-430-201-970	Sequence 970, App
C 22	20	1.7	507	17 US-10-430-201-971	Sequence 971, App
C 23	20	1.7	596	17 US-10-767-701-5378	Sequence 5378, Ap
C 24	20	1.7	848	13 US-10-425-114-20418	Sequence 20418, A
C 25	20	1.7	863	13 US-10-027-632-150405	Sequence 150405, A
C 26	20	1.7	863	16 US-10-027-632-150405	Sequence 150405, A
C 27	20	1.7	1192	16 US-10-120-988-90	Sequence 90, Appl
C 28	20	1.7	1255	13 US-10-425-114-19266	Sequence 19266, A
C 29	20	1.7	1290	15 US-10-146-733-16	Sequence 16, Appl
C 30	20	1.7	1383	17 US-10-437-963-37464	Sequence 37464, A
C 31	20	1.7	1526	13 US-10-425-114-31002	Sequence 31002, A
C 32	20	1.7	1576	13 US-09-852-386-87	Sequence 87, Appl
C 33	20	1.7	1737	15 US-10-156-761-3579	Sequence 3579, Ap
C 34	20	1.7	1814	15 US-10-060-036-182	Sequence 182, App
C 35	20	1.7	1910	9 US-09-789-836-1	Sequence 1, Appl
C 36	20	1.7	1910	10 US-09-789-831-1	Sequence 1, Appl
C 37	20	1.7	1910	17 US-10-415-325-18	Sequence 18, Appl
C 38	20	1.7	1943	15 US-10-146-733-14	Sequence 14, Appl
C 39	20	1.7	2080	13 US-10-257-022-30	Sequence 30, Appl
C 40	20	1.7	2504	14 US-10-108-605-246	Sequence 246, App
C 41	20	1.7	2616	17 US-10-437-963-75800	Sequence 75800, A
C 42	20	1.7	2874	9 US-09-801-260-3	Sequence 3, Appl
C 43	20	1.7	2874	15 US-10-145-586-42	Sequence 42, Appl
C 44	20	1.7	3112	16 US-10-120-988-261	Sequence 261, App
C 45	20	1.7	3226	9 US-09-801-260-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-10-104-047-1104
; Sequence 1104, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cdna
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1104
; LENGTH: 2051
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1104

Query Match	83.6%	Score	976;	DB	16;	Length	2051;
Best Local Similarity	99.8%	Pred. No.	0;				
Matches	1076;	Conservative	0;	Mismatches	2;	Indels	0;
Gaps	0;						
QY	91	AGTCCTGAAGGGCTCCAGGCTCGCTTCACTGACCGCTCCAGGGCTGGAGCTCAT	150				
Db	250	AGTCCTGAAGGGCTCCAGGCTCGCTTCACTGACCGCTCCAGGGCTGGAGCTCAT	309				
QY	151	CATGTGGGCTCTCAGTGAATGGTGTGTAAAGCGTCAGGCCCATGAGCCCATCATCAC	210				
Db	310	CATGTGGGCTCTCAGTGAATGGTGTGTAAAGCGTCAGGCCCATGAGCCCATCATCAC	369				
QY	211	CAATGACCGCTTCACTCTTCAAGAGTACGACGAGGCGGGAACCTTCACTCGGAGATGAT	270				
Db	370	CAATGACCGCTTCACTCTTCAAGAGTACGACGAGGCGGGAACCTTCACTCGGAGATGAT	429				
QY	271	CATCCAAATGTGGAGCCCGCTGATTCGGGGGAACATCAGATGAGCCTCCGAACAGTCG	330				


```

RESULT 5
US-10-437-963-81852/c
; Sequence 81852, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kowalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 81852
; LENGTH: 1284
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_81337C.1
US-10-437-963-81852

Query Match          2.0%; Score 23; DB 17; Length 1284;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

US-10-437-963-81852/C
; Sequence 81852, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbaruk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 81852
; LENGTH: 1284
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_81337C.1
US-10-437-963-81852

Query Match          2.0%; Score 23; DB 17; Length 1284;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 800 CGCTGCTGCTGCTGCGCGCGTCG 822
Db 156 CGCTGCTGCTGCTGCGCGCGTCG 134

RESULT 6

US-09-770-445-930
; Sequence 930, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2023US (PARA-012PRV)
; CURRENT APPLICATION NUMBER: US/09/770,445
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 930
; LENGTH: 767
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-445-930

Query Match 1.9%; Score 22; DB 9; Length 767;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 814 CGCGCGCTGCTGTTGTGCGTGC 835
Db 672 CGCGCGCTGCTGTTGTGCGTGC 693

RESULT 7

US-09-220-920-65
; Sequence 65, Application US/09220920
; Patent No. US2002002269A1
; GENERAL INFORMATION:
; APPLICANT: Milbrandt, Jeffrey D.
; APPLICANT: Baloh, Robert H.
; TITLE OF INVENTION: Artemin, A No. US2002002269A1 Neurotrophic Factor
; FILE REFERENCE: 6029-7996
; CURRENT APPLICATION NUMBER: US/09/220,920
; CURRENT FILING DATE: 1998-12-24
; EARLIER APPLICATION NUMBER: 09/163,283
; EARLIER FILING DATE: 1998-09-29
; EARLIER APPLICATION NUMBER: 60/108,148
; EARLIER FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/218,698
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 65
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-920-65

Query Match 1.9%; Score 22; DB 9; Length 1203;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCTGCGCGCGTCG 822
Db 51 GCTGCTGCTGCTGCGCGCGTCG 72

RESULT 8

US-09-828-366-15
; Sequence 15, Application US/09828366
; Patent No. US20020010137A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Klein, Robert D.
; APPLICANT: Napier, Mary
; APPLICANT: Wood, William I.
; APPLICANT: Yuan, Jean
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC
; TITLE OF INVENTION: CELL GROWTH
; FILE REFERENCE: P1694R1C1
; CURRENT APPLICATION NUMBER: US/09/828,366
; CURRENT FILING DATE: 2001-04-05
; Prior filing data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 29
; SEQ ID NO 15
; LENGTH: 1829
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-828-366-15

Query Match 1.9%; Score 22; DB 9; Length 1829;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCTGCGCGCGTCG 822
Db 88 GCTGCTGCTGCTGCGCGCGTCG 109

RESULT 9

US-10-152-319A-1498
; Sequence 1498, Application US/10152319A
; Publication No. US20040072160A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castie, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10

```
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1498
; LENGTH: 4142
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_012544
US-10-152-319A-1498

Query Match      1.9%; Score 22; DB 12; Length 4142;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      801 GCTGCTGCTGCTGCCGCGTCG 822
Db      95 GCTGCTGCTGCTGCCGCGTCG 116

RESULT 10
US-10-437-963-41506/c
; Sequence 41506, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 41506
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_44844C.1
US-10-437-963-41506

Query Match      1.8%; Score 21; DB 17; Length 462;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      802 CTGCTGCTGCTGCCGCGTCG 822
Db      133 CTGCTGCTGCTGCCGCGTCG 113

RESULT 11
US-09-918-995-11051/c
; Sequence 11051, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
```

```
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11051
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(493)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-11051

Query Match      1.8%; Score 21; DB 10; Length 493;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      804 GCTGCTGCTGCCGCGTCGTT 824
Db      443 GCTGCTGCTGCCGCGTCGTT 423

RESULT 12
US-10-255-536-128/c
; Sequence 128, Application US/10255536
; Publication No. US20030087807A1
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Ralph J.
; TITLE OF INVENTION: Methods for Identifying Compounds for
; TITLE OF INVENTION: Motion Sickness, Vertigo and Other Disorders Related to
; TITLE OF INVENTION: Balance and the Perception of Gravity
; FILE REFERENCE: P-NI 3864
; CURRENT APPLICATION NUMBER: US/10/255,536
; CURRENT FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: US/09/669,751
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 60/168,579
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 128
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Drosophila
; APPLICANT: Drosophila
US-10-255-536-128

Query Match      1.8%; Score 21; DB 15; Length 550;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      837 ACTGCTGCTGCCGTTGTTGTT 857
Db      317 ACTGCTGCTGCCGTTGTTGTT 297

RESULT 13
US-10-767-701-28842/c
; Sequence 28842, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
```

; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 28842
; LENGTH: 602
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 8088472
US-10-767-701-28842

Query Match 1.8%; Score 21; DB 17; Length 602;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 803 TGTCTGCTGCGCGCGTGT 823
Db 206 TGTCTGCTGCGCGCGTGT 186

RESULT 14

US-09-930-213-20/c
; Sequence 20, Application US/09930213
; Publication No. US20030170625A1
; GENERAL INFORMATION:
; APPLICANT: ROSENTHAL, ANDRE
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHAFER, REINHARD
; APPLICANT: ZUBER, JOHANNES
; APPLICANT: TCHE-NITSE, OLEG
; APPLICANT: GRIPS, MARTIN
; APPLICANT: HELLMER, MARTIN
; APPLICANT: SCHMITZ, ANNE-CHANTAL
; APPLICANT: SERS, CHRISTINE
; TITLE OF INVENTION: DETECTION OF DIFFERENTIAL GENE EXPRESSIONS
; FILE REFERENCE: ALBRE-14
; CURRENT APPLICATION NUMBER: US/09/930,213
; CURRENT FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: DE 10004102.7
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 885
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 1558
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (16)
; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (427)
; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (492)
; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (551)
; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (728)
; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (744)
; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (979)

; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1146)
; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1243)
; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1343)
; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1346)
; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1352)
; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1423)
; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1535)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-09-930-213-20

Query Match 1.8%; Score 21; DB 10; Length 1558;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCTGCGCGTGC 821
Db 223 GCTGCTGCTGCTGCGCGTGC 203

RESULT 15

US-10-087-192-185/c
; Sequence 185, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 185
; LENGTH: 3362
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-087-192-185

Query Match 1.8%; Score 21; DB 13; Length 3362;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCTGCGCGTGC 821
Db 947 GCTGCTGCTGCTGCGCGTGC 927

Search completed: September 15, 2004, 07:20:16
Job time : 620.673 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 20:48:25 ; Search time 5096.01 Seconds
(without alignments)
10546.560 Million cell updates/sec

Title: US-09-729-264-5

Perfect score: 1240

Sequence: 1 aggtgtgagtcgcccgaaca.....gtaatacaactgtagtatag 1240

Scoring table: OLIGO NUC

Gapop_60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6784142

Minimum DB seq length: 16

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_man.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1240	100.0	1240	6	AX380400	Sequence
2	1125	90.7	2051	6	AX747579	Sequence
3	1125	90.7	2051	6	AK092516	Homo sapi
4	1074	86.6	1175	6	AX380396	Sequence
5	976	78.7	1168	6	AX380398	Sequence
6	739	59.6	1139	6	AX380402	Sequence
7	302	24.4	142742	9	AF121782	Homo sapi
8	302	24.4	340000	9	HS21C080	Homo sapi
9	200	16.1	182532	9	CH179K04	Pan trogl
10	145	11.7	156288	9	BS000162	Pan trogl
11	119	9.6	40205	9	AF045450	Homo sapi
12	119	9.6	170121	9	AF064860	Homo sapi
13	100	8.1	199665	9	AF064857	Homo sapi
14	72	5.8	192219	9	RP43002119	Pan trogl
15	25	2.0	6470	3	AB090820	Anopheles
16	24	1.9	5178	3	AY119603	Drosophila
17	24	1.9	8152	3	DMTOC	Human DNA
18	24	1.9	55256	9	HSJ735G18	Human DNA
19	24	1.9	59693	2	AC099814	Homo sapi
20	24	1.9	77137	3	AC004422	Drosophila
21	24	1.9	83876	2	AC020009	Drosophila
22	24	1.9	94296	2	AL160261	Homo sapi
23	24	1.9	111645	9	AL138767	Human DNA
24	24	1.9	136746	2	AC006191	Homo sapi
25	24	1.9	148102	8	AP003328	Oryza sat
26	24	1.9	148762	8	AP002843	Oryza sat
27	24	1.9	161529	9	AP001998	Homo sapi
28	24	1.9	169931	3	AC008321	Drosophila
29	24	1.9	171400	9	AP006437	Homo sapi
30	24	1.9	181595	9	AL139350	Human DNA
31	24	1.9	184554	3	AC099022	Drosophila
32	24	1.9	196857	9	AC027239	Homo sapi
33	24	1.9	206773	2	AC010932	Homo sapi
34	24	1.9	268984	3	AE001274	Leishmani
35	24	1.9	314957	3	AE003581	Drosophila
36	24	1.9	342614	2	AC108705	Homo sapi
37	23	1.9	1155	10	AFS37215	Mus muscu
38	23	1.9	1195	6	AX380404	Sequence
39	23	1.9	29865	2	AC020047	Drosophila
40	23	1.9	100494	9	HSJ514B11	Human DNA
41	23	1.9	106256	3	AC108135	Leishmani
42	23	1.9	124612	2	AC105442	Leishmani
43	23	1.9	138764	9	AC006387	Homo sapi
44	23	1.9	144500	2	AC010231	Homo sapi
45	23	1.9	149964	2	AC120145	Mus muscu

ALIGNMENTS

RESULT 1
AX380400
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AX380400
Sequence 5 from Patent WO200710.
AX380400
AX380400.1 GI:19575330
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Welcher,A.A., Sarmiento,U.M., Schultz,H.J. and Chute,H.T.
B7-like molecules and uses thereof
Patent: WO 0200710-A 5 03-JAN-2002;

AX380400
Sequence 5 from Patent WO200710.
AX380400
AX380400.1 GI:19575330
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Welcher,A.A., Sarmiento,U.M., Schultz,H.J. and Chute,H.T.
B7-like molecules and uses thereof
Patent: WO 0200710-A 5 03-JAN-2002;

Pred. No. is the number of results predicted by chance to have a

Amgen, Inc. (US)		Location/Qualifiers	
1..1240		/organism="Homo sapiens"	
/mol_type="unassigned DNA"		/db_xref="taxon:9606"	
80..1240		/note="unnamed protein product"	
/codon_start=1		/protein_id="CAD28218.1"	
/db_xref="GI:19575331"		/db_xref="REMTREMBL:CAD28218"	
/translation="WERHLITVPEAVSGSGNEVLEGQONATVLKGSQARFNCTVSGG WKLIMWALSDMVLSVRPMEPIITNDRFTSORYDOGNETSEMIHNVPSPDSGNIRC SLQNSRLHSAVLTWQVMEGLFIPSVNLVAENEPCEVTLPSHWIRLPDIISWELGLL VSHSYTFVPEPSDLQASVILALTFQSNGLTLCVATWKSARKASATVNLVLRCPQ DTGGGINIFGVLSLSPLGFLPTWKGVLGLAGTMLLPTCLTIRCCRRRCRCCG NCCRCPCFCRKRGRFIOFKKSEKTKETETESNGNSYNSDEQKTTETASLP PKSCSSDPEQRNSSCGPPHQADQRP RP PASHPQASFNLASPEKVSNTTVV"			
ORIGIN			
Query Match		100.0%; Score 1240; DB 6; Length 1240;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 1240; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	AGGTGTGAGTCCAGCCAAACAGTGTGGATCAGTTTCTTAGGCTGCCATAACAAAGCACCAT	60
DB	1	AGGTGTGAGTCCAGCCAAACAGTGTGGATCAGTTTCTTAGGCTGCCATAACAAAGCACCAT	60
QY	61	AACCTGGTGGCTTAGAACAAATGAAAGGCAATTTGCTCAOAGTTTCCAGAGCTGTAGTTTC	120
DB	61	AACCTGGTGGCTTAGAACAAATGAAAGGCAATTTGCTCAOAGTTTCCAGAGCTGTAGTTTC	120
QY	121	TGGGTCTGTTAATGAAGTCTAGAACGCCCCCGAGATGCAACAGTCTGAGAGGCTCCCA	180
DB	121	TGGGTCTGTTAATGAAGTCTAGAACGCCCCCGAGATGCAACAGTCTGAGAGGCTCCCA	180
QY	181	GGCTCGCTTCAACTGACACCGCTCTCCAGGCTGGAAGCTCATATGTTGGCTCTCAGTGA	240
DB	181	GGCTCGCTTCAACTGACACCGCTCTCCAGGCTGGAAGCTCATATGTTGGCTCTCAGTGA	240
QY	241	CATGGTGGTCTAAGCGTCAAGGCCATGGAGCCCATCATCACCATAACCGCTTACCTC	300
DB	241	CATGGTGGTCTAAGCGTCAAGGCCATGGAGCCCATCATCACCATAACCGCTTACCTC	300
QY	301	TCAGAGTACGACACAGGCGGGAACTTCACTCGGAGATGATCATCCAAATGTGGAGCC	360
DB	301	TCAGAGTACGACACAGGCGGGAACTTCACTCGGAGATGATCATCCAAATGTGGAGCC	360
QY	361	CAGTGATTCGGGGAACATCAGATGCAGCTCCAGAACAGTCGCCTGCATGGATCTGCTTA	420
DB	361	CAGTGATTCGGGGAACATCAGATGCAGCTCCAGAACAGTCGCCTGCATGGATCTGCTTA	420
QY	421	CTTTACCGTCCAAAGTTATGGAGAGCTGTTCATTTCCAGTGTGTTAATCTTGTAGTCGTGA	480
DB	421	CTTTACCGTCCAAAGTTATGGAGAGCTGTTCATTTCCAGTGTGTTAATCTTGTAGTCGTGA	480
QY	481	GAATGAACTTGTGAAGTACTTGTCTACCTCCTCACACTGGACCCGCTCCCGGATATTTTC	540
DB	481	GAATGAACTTGTGAAGTACTTGTCTACCTCCTCACACTGGACCCGCTCCCGGATATTTTC	540
QY	541	CTGGAGCTCGGTCTCTCGTTCAGCCATTAAGCTATTATTTTGTTCGGAGCCCGAGCGA	600
DB	541	CTGGAGCTCGGTCTCTCGTTCAGCCATTAAGCTATTATTTTGTTCGGAGCCCGAGCGA	600
QY	601	CTTTCAAAGTGCAGTGAGCATCTCTGGCTCTGACCCACAGAGCAATGGGACTTTGACTTG	660
DB	601	CTTTCAAAGTGCAGTGAGCATCTCTGGCTCTGACCCACAGAGCAATGGGACTTTGACTTG	660
QY	661	CGTGGCTACCTGGAGAGGCTGGAAGGCCCGCAAGTCTGCAACTGTGTAATCTCACTGTGAT	720
DB	661	CGTGGCTACCTGGAGAGGCTGGAAGGCCCGCAAGTCTGCAACTGTGTAATCTCACTGTGAT	720

QY	721	TCGGTGTCCCAAGACACTGGAGGTGGTATTAATATTCCAGGTGTATTAATCAAGTTTACC	780
DB	721	TCGGTGTCCCAAGACACTGGAGGTGGTATTAATATTCCAGGTGTATTAATCAAGTTTACC	780
QY	781	GAGTTTAGGTTTTTTCATTCGCTTACCTGGGCAAAAGTTGGACTTGGAGTGGAGCACCAT	840
DB	781	GAGTTTAGGTTTTTTCATTCGCTTACCTGGGCAAAAGTTGGACTTGGAGTGGAGCACCAT	840
QY	841	GCTTCTGACGCGCAGCTGTACTCTTACAATAGCTGCTGCTGCGCGCTGTTGTTG	900
DB	841	GCTTCTGACGCGCAGCTGTACTCTTACAATAGCTGCTGCTGCGCGCTGTTGTTG	900
QY	901	TGGCTGCAACTGCTGCTGCGCTTGTGTTTCTGCTGTAGAAAGAAAAAGAGGATTTTCGTAT	960
DB	901	TGGCTGCAACTGCTGCTGCGCTTGTGTTTCTGCTGTAGAAAGAAAAAGAGGATTTTCGTAT	960
QY	961	TCAATTTCAAAGAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGG	1020
DB	961	TCAATTTCAAAGAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGG	1020
QY	1021	AAATGAAAACCTCGGCTCAAAATTCAGATGAACAAAAGACCAAGAAACCGCTTCTCTCCC	1080
DB	1021	AAATGAAAACCTCGGCTCAAAATTCAGATGAACAAAAGACCAAGAAACCGCTTCTCTCCC	1080
QY	1081	TCCCAATCTGTGAATCCAGTGTCTTGAACAAAGAAACAGTAGTGTGGCCCTCTCTCA	1140
DB	1081	TCCCAATCTGTGAATCCAGTGTCTTGAACAAAGAAACAGTAGTGTGGCCCTCTCTCA	1140
QY	1141	CMAGCGGCTGATCAACGTCACCCAGGCGCAGCAAGTCATCCACAGGCTTCTTTTAATCT	1200
DB	1141	CMAGCGGCTGATCAACGTCACCCAGGCGCAGCAAGTCATCCACAGGCTTCTTTTAATCT	1200
QY	1201	GGCCAGTCTCTGAGAGGTCAGTAATACAACTGTAGTATAG	1240
DB	1201	GGCCAGTCTCTGAGAGGTCAGTAATACAACTGTAGTATAG	1240

RESULT 2

AX747579

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..2051

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

ORIGIN

Query Match

Best Local Similarity

Matches 1125; Conservative

90.7%; Score 1125; DB 6; Length 2051;

100.0%; Pred. No. 0;

0; Mismatches 0; Indels 0; Gaps 0;

QY

116

GGTTCTGGGTCTGGTAATGAAGTCATAGAAGGCCCGCCAGATGCAACAGTCTCTGAAGGC

175

DB

203

GGTTCTGGGTCTGGTAATGAAGTCATAGAAGGCCCGCCAGATGCAACAGTCTCTGAAGGC

262

QY

176

TCCAGGCTCGCTTCAACTGACCGCTCTCCAGGCTGGAAGCTCATCATGTGGGCTCTC

235

```

Db 263 TCCAGAGCTCGCTTCAACTGCAACGCTCTCCAGGCGTGAAGCTCATATGTTGGGCTCTC 322
QY 236 AGTGACATGGTGGTCTAAGCGCTCAGGGCCATGAGGCCATCATCACCAGTCAACGCTTC 295
Db 323 AGTGACATGGTGGTCTAAGCGCTCAGGGCCATGAGGCCATCATCACCAGTCAACGCTTC 382
QY 296 ACCTCTCAGAGTACGACAGGCGGGAACTTCACTCGGAGATGATCATCCACAATGTG 355
Db 383 ACCTCTCAGAGTACGACAGGCGGGAACTTCACTCGGAGATGATCATCCACAATGTG 442
QY 356 GAGCCAGTGAATCGGGGAACATCAGATGCGAGCTCCAGAACAGTGGCTGATGGATCT 415
Db 443 GAGCCAGTGAATCGGGGAACATCAGATGCGAGCTCCAGAACAGTGGCTGATGGATCT 502
QY 416 GCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTTCATTCACAGTGTAACTTGTAGTC 475
Db 503 GCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTTCATTCACAGTGTAACTTGTAGTC 562
QY 476 GCTGAGATGAACCTTGTGAAGTTACTTGTCTACCCCTCACACTGGACCGGCTCCCGAT 535
Db 563 GCTGAGATGAACCTTGTGAAGTTACTTGTCTACCCCTCACACTGGACCGGCTCCCGAT 622
QY 536 ATTTCTGGAGCTCGTCTCTCGTCAAGCAATTCAGCTATTATTTGTTCCGAGGCC 595
Db 623 ATTTCTGGAGCTCGTCTCTCGTCAAGCAATTCAGCTATTATTTGTTCCGAGGCC 682
QY 596 AGGACCTTCAAGTGCAGTGAGCATCTCGTCTCGACCCACAGAGCAATGGACATTG 655
Db 683 AGGACCTTCAAGTGCAGTGAGCATCTCGTCTCGACCCACAGAGCAATGGACATTG 742
QY 656 ACTTGGCTGGTACTGAGAGAGCTGAAGCCCGCAAGTGTGCAACTGTAATCTCACT 715
Db 743 ACTTGGCTGGTACTGAGAGAGCTGAAGCCCGCAAGTGTGCAACTGTAATCTCACT 802
QY 716 GTGATTCGGTGTCCCAAGACACTGGAGTGTATTAATATTCAGTGTATTAAGT 775
Db 803 GTGATTCGGTGTCCCAAGACACTGGAGTGTATTAATATTCAGTGTATTAAGT 862
QY 776 TTACCGAGTTAGTGTCTTCAATTCGCTACTTGGGCAAGTGTGACTTGGACTAGCAGGC 835
Db 863 TTACCGAGTTAGTGTCTTCAATTCGCTACTTGGGCAAGTGTGACTTGGACTAGCAGGC 922
QY 836 ACCATGCTTCTGAGCGCGAGTGTACTCTTACAAATACGCTGCTCTCGCCGCGTCT 895
Db 923 ACCATGCTTCTGAGCGCGAGTGTACTCTTACAAATACGCTGCTCTCGCCGCGTCT 982
QY 896 TGTGTGGCTGCAACTGCTGCTCGTGTGTTGTTCTGCTGTAGAGAAAGAGGATTT 955
Db 983 TGTGTGGCTGCAACTGCTGCTCGTGTGTTGTTCTGCTGTAGAGAAAGAGGATTT 1042
QY 956 CGTATTCAATTTCAAAGAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAGAA 1015
Db 1043 CGTATTCAATTTCAAAGAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAGAA 1102
QY 1016 AGTGGAAATGAANAATCCCGGCTCAATTCAGATGAAACAAAGACACAGAAACCGCTTCT 1075
Db 1103 AGTGGAAATGAANAATCCCGGCTCAATTCAGATGAAACAAAGACACAGAAACCGCTTCT 1162
QY 1076 CTCCCTCCCAATTCCTGTGAATCCAGTGATCCCTGAACAAAGAAACAGTACTGTGGCCCT 1135
Db 1163 CTCCCTCCCAATTCCTGTGAATCCAGTGATCCCTGAACAAAGAAACAGTACTGTGGCCCT 1222
QY 1136 CTTACACAGCGGCTGATCAACGTTCCACCGCCAGCAGTCAATCCACAGGCTTCTTTT 1195
Db 1223 CTTACACAGCGGCTGATCAACGTTCCACCGCCAGCAGTCAATCCACAGGCTTCTTTT 1282
QY 1196 AATCTGGCAGCTCTGAGAGGTCAGTAATACAACTGTAGTATAG 1240
Db 1283 AATCTGGCAGCTCTGAGAGGTCAGTAATACAACTGTAGTATAG 1327

```

RESULT 3

```

AK092516
LOCUS Homo sapiens cDNA FLJ35197 fis, clone PLACE6017788, highly similar
DEFINITION to IGSP5.
ACCESSION AK092516
VERSION GI:21751130
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
Oshima, A., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K.
and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2051)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
Location/Qualifiers
1..2051
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE6017788"
/tissue_type="placenta"
/clone_lib="PLACE6"
/note="Cloning vector: pME18SFL3"
ORIGIN
Query Match 90.7%; Score 1125; DB 9; Length 2051;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 116 GGTTCCTGGGCTGGTAAATGAAGTCATAGAAGGCCCCAGAAATCAACAGTCTCTGAAGGC 175
Db 203 GGTTCCTGGGCTGGTAAATGAAGTCATAGAAGGCCCCAGAAATCAACAGTCTCTGAAGGC 262
QY 176 TCCAGGCTCGCTTCAACTGCACCGTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTC 235
Db 263 TCCAGGCTCGCTTCAACTGCACCGTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTC 322
QY 236 AGTGACATGGTGGTGTAAAGCGTCAAGCCCATGGAGCCCATCATCAACATACCGCTTC 295
Db 323 AGTGACATGGTGGTGTAAAGCGTCAAGCCCATGGAGCCCATCATCAACATACCGCTTC 382
QY 296 ACCTCTCAGAGTACGACAGGCGGGAACTTCACTCGGAGATGATCATCCACAATGTG 355
Db 383 ACCTCTCAGAGTACGACAGGCGGGAACTTCACTCGGAGATGATCATCCACAATGTG 442
QY 356 GAGCCAGTGAATTCGGGGAAACATCAGATGCGAGCTCCAGAACAGTGGCTGATGGATCT 415
Db 443 GAGCCAGTGAATTCGGGGAAACATCAGATGCGAGCTCCAGAACAGTGGCTGATGGATCT 502
QY 416 GCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTTCATTCACAGTGTAACTTGTAGTC 475
Db 503 GCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTTCATTCACAGTGTAACTTGTAGTC 562

```

QY	476	GCTGAGAAAGAACTTGTGAAGTACTTGTCTACCTCACACTGACCGCGTCCCGAT	535	/mol_type="unassigned DNA"
Db	563	GCTGAGAAAGAACTTGTGAAGTACTTGTCTACCTCACACTGACCGCGTCCCGAT	622	/db_xref="taxon:9606"
QY	536	ATTTCCTGGAGCTCGGTCTCTGTCAGCAATCAAGCTATTATTTTGTTCGGAGGCC	595	27..1175
Db	623	ATTTCCTGGAGCTCGGTCTCTGTCAGCAATCAAGCTATTATTTTGTTCGGAGGCC	682	/note="unnamed protein product"
QY	596	AGCGACCTTCAAAGTCAGTGAAGTCTCTGCTGACCCACAGAGCAATGGACTTTG	655	/codon_start=1
Db	683	AGCGACCTTCAAAGTCAGTGAAGTCTCTGCTGACCCACAGAGCAATGGACTTTG	742	/protein_id="CAD28216.1"
QY	656	ACTTGGTGGTCACTTGAAGAGCTGAAGGCCCGCAAGTCTGCACTGTAATCTCACT	715	/db_xref="GI:19575327"
Db	743	ACTTGGTGGTCACTTGAAGAGCTGAAGGCCCGCAAGTCTGCACTGTAATCTCACT	802	/translation="MGLVIFLHSGSGNEVIEGPQNAVTLKGSQARFNCTVSGWKLI
QY	716	GTGATTCGGTGTCCCAAGACACTGAGGTGGTATTATATATCCAGGTGATTATCAAGT	775	/db_xref="REMTREMBL:CAD28216"
Db	803	GTGATTCGGTGTCCCAAGACACTGAGGTGGTATTATATATCCAGGTGATTATCAAGT	862	MMALSDMVLSVRPMEPIITNDRFTSGSRYDQGNFTSBMIHNVPSPSGNTRCSLON
QY	776	TTACCGAGTTTAGGTTTTTCAATGCTACTTGGGGCAAAAGTTGGACTTAGCAGGC	835	SRLLHSAVLTQVMGELFIPSNLVVAENPEVETCLPSHWTRLPDITWIRCPQDTGHS
Db	863	TTACCGAGTTTAGGTTTTTCAATGCTACTTGGGGCAAAAGTTGGACTTAGCAGGC	922	SYTFPEPSDILQSAVSIILALTPQSNGLTICVATKSLKARKSATVNLTVIRCPQDTGHS
QY	836	ACCATGCTTCTGAGCGGAGCGTGTACTCTTACAATACGCTGCTGTCTGCGCGCTCGT	895	GINTPGVLSLPSUGFSIPTWKGVLGLAGTLMLLTPTCTLTATCCCCRRCCGCGNCCC
Db	923	ACCATGCTTCTGAGCGGAGCGTGTACTCTTACAATACGCTGCTGTCTGCGCGCTCGT	982	RCFCRRKRGFRIOFKSEKKTNETETESNGENSINSDEKTTDTASLPKSC
QY	896	TGTTGTGGTGTCAACTGCTGCTGCCGTGTTGTTTCTGTGTAGAGAAAGAGGATTT	955	BSSDDEQRNSCGPPHQRDQRPFRPASHPQASFNLASPERKVSNTTVV
Db	983	TGTTGTGGTGTCAACTGCTGCTGCCGTGTTGTTTCTGTGTAGAGAAAGAGGATTT	1042	
QY	956	CGTATTCAATTTCAAAGAAATCTGAAAAAGAGAGACAAACAAAGAACTGAGACAGAA	1015	
Db	1043	CGTATTCAATTTCAAAGAAATCTGAAAAAGAGAGACAAACAAAGAACTGAGACAGAA	1102	
QY	1016	AGTGGAAATGAAGTCCGGCTCAATTCAGATGACAAAGAACCCAGAGAAACCGCTTCT	1075	
Db	1103	AGTGGAAATGAAGTCCGGCTCAATTCAGATGACAAAGAACCCAGAGAAACCGCTTCT	1162	
QY	1076	CTCCCTCCCAATCCTGTGAATCAGTGTATCTGAAACAAAGAAACAGTAGCTGTGGCCCT	1135	
Db	1163	CTCCCTCCCAATCCTGTGAATCAGTGTATCTGAAACAAAGAAACAGTAGCTGTGGCCCT	1222	
QY	1136	CCTCACCAGCGGGTGATCAACAGTCCAGGCCAGCAAGTATCCAGGCTTCTTTT	1195	
Db	1223	CCTCACCAGCGGGTGATCAACAGTCCAGGCCAGCAAGTATCCAGGCTTCTTTT	1282	
QY	1196	AATCTGGCCAGTCTCGAGAGGTCAAGTAAATACAACTGTAGTATAG	1240	
Db	1283	AATCTGGCCAGTCTCGAGAGGTCAAGTAAATACAACTGTAGTATAG	1327	
RESULT 4				
AX380396				
LOCUS				
DEFINITION				Sequence 1 from Patent WO0200710.
ACCESSION				AX380396
VERSION				AX380396.1 GI:19575326
KEYWORDS				
SOURCE				Homo sapiens (human)
ORGANISM				Homo sapiens
				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
				Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE				1
AUTHORS				Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
TITLE				B7-like molecules and uses thereof
JOURNAL				Patent: WO 0200710-A 1 03-JAN-2002;
FEATURES				Amgen, Inc. (US)
				Location/Qualifiers
				1..1175
				/organism="Homo sapiens"

Query Match

Best Local Similarity

Matches 1124; Conservative

86.6%; Score 1074; DB 6; Length 1175;

Pred. No. 0;

1; Indels 0; Gaps 0;

0; Mismatches

QY	116	GGTTCTGGGTCTGGTAATGAAGTCATAGAAGGCCCCAGAAATGCAACAGTCTCGAAGGCC	175
Db	51	GGTTCTGGGTCTGGTAATGAAGTCATAGAAGGCCCCAGAAATGCAACAGTCTCGAAGGCC	110
QY	176	TCCAGGCTCGCTTCAATGCAACGCTCTCCAGGGGTGGAAGCTCATCATGTGGGCTCTC	235
Db	111	TCCAGGCTCGCTTCAATGCAACGCTCTCCAGGGGTGGAAGCTCATCATGTGGGCTCTC	170
QY	236	AGTGACATGGTGTGCTAGCGTCAGGCCCATGGAGCCCATCATCAATGACCGCTTC	295
Db	171	AGTGACATGGTGTGCTAGCGTCAGGCCCATGGAGCCCATCATCAATGACCGCTTC	230
QY	296	ACCTCTCAGAGGTACACAGCGCGGAACTTCACTCGGAGATGATCATCAATGTG	355
Db	231	ACCTCTCAGAGGTACACAGCGCGGAACTTCACTCGGAGATGATCATCAATGTG	290
QY	356	GAGCCCAAGTATTCGGGGAAACATCAGATGAGCTCCAGAACAGTGCCTGCGATGATCT	415
Db	291	GAGCCCAAGTATTCGGGGAAACATCAGATGAGCTCCAGAACAGTGCCTGCGATGATCT	350
QY	416	GCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCATTCCTCAGTGTATCTGTAGTC	475
Db	351	GCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCATTCCTCAGTGTATCTGTAGTC	410
QY	476	GCTGAGAAATGAACCTTGTGAAGTTACTTGTCTACCTCACAAGTCCCGGCTCCCGAT	535
Db	411	GCTGAGAAATGAACCTTGTGAAGTTACTTGTCTACCTCACAAGTCCCGGCTCCCGAT	470
QY	536	ATTTCCTGGAGCTCGGTCTCTGGTCAGGCATTCAGCTATTATTTTGTTCGGAGGCC	595
Db	471	ATTTCCTGGAGCTCGGTCTCTGGTCAGGCATTCAGCTATTATTTTGTTCGGAGGCC	530
QY	596	AGCGACCTTCAAAGTCAGTGAAGTCTCTGCTGACCCACAGAGCAATGGAGCTTTG	655
Db	531	AGCGACCTTCAAAGTCAGTGAAGTCTCTGCTGACCCACAGAGCAATGGAGCTTTG	590
QY	656	ACTTGGTGGTCACTCTGGAAGAGCTTGAAGGCCCGCAAGTCTGCAACTGTAATCTCACT	715
Db	591	ACTTGGTGGTCACTCTGGAAGAGCTTGAAGGCCCGCAAGTCTGCAACTGTAATCTCACT	650
QY	716	GTGATTCGGTGTCCCAAGACACTGAGGTGGTATTATATATCCAGGTGATTATCAAGT	775
Db	651	GTGATTCGGTGTCCCAAGACACTGAGGTGGTATTATATATCCAGGTGATTATCAAGT	710
QY	776	TTACCGAGTTTAGGTTTTTCAATGCTACTTGGGGCAAAAGTTGGACTTAGCAGGC	835
Db	711	TTACCGAGTTTAGGTTTTTCAATGCTACTTGGGGCAAAAGTTGGACTTAGCAGGC	770
QY	836	ACCATGCTTCTGAGCGGAGCGTGTACTCTTACAATACGCTGCTGTCTGCGCGCTCGT	895
Db	771	ACCATGCTTCTGAGCGGAGCGTGTACTCTTACAATACGCTGCTGTCTGCGCGCTCGT	830

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
TITLE B7-like molecules and uses thereof
JOURNAL Patent: WO 020710-A 7 03-JAN-2002;
Amgen, Inc. (US)
FEATURES
source Location/Qualifiers
1..1139
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..1134
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD28219.1"
/db_xref="GI:19575333"
/db_xref="EMBL:CAD28219"
/translation="MVAGAMENRDPGSGNEVIEGPONARVLKSGQARFNCTVSQG
WKLIMWLSDMVLVSVRPMEPIITNDRFTSORYDOGNLTSEMIHNVPSPSGNIR
SLONSLHGSAYLTQVWVGEPIFIPSNLVAENPECEVCLPSHWRLPDIISWELGL
VSHSYTFVPEPSDLSQASVILALTPQSNGLTTCVATWKLKARKSATVNLVIRCPQ
DTGGNINPGLSSLPSPGSLFTWGVKVLGAGNMLTPTCTLTIRCCCCRRRCGC
NCCRCFCRCRKGRLKRRQTKLRQKRVKMTPATIQNKRPTQPTLLSLNPVNPV
ILNKETVAALLTSLGLNVHPGQVTHRELLIWPVLRSSVQL"
ORIGIN
Query Match 59.6%; Score 739; DB 6; Length 1139;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 789; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 163 AGTCCTGAAGGGCTCCAGGCTCGCTCACTCACTGCTCCAGGGCTGGAAGTCTCAT 222
DB 84 AGTCCTGAAGGGCTCCAGGCTCGCTCACTCACTGCTCCAGGGCTGGAAGTCTCAT 143
QY 223 CATGTGGGCTCTCAGTGACATGTGTGTGCTAAAGCGTCAGGCCCATGAGCCCATCATCAC 282
DB 144 CATGTGGGCTCTCAGTGACATGTGTGTGCTAAAGCGTCAGGCCCATGAGCCCATCATCAC 203
QY 283 CAATGACCGGTTCACTCTCAGAGGTACGACCGAGGGGGGAACCTTCACTCGGAGATGAT 342
DB 204 CAATGACCGGTTCACTCTCAGAGGTACGACCGAGGGGGGAACCTTCACTCGGAGATGAT 263
QY 343 CATCCCAATGTGGAGCCAGTATTCTGGGGGAACATCAGATGACGCTCCAGAACAGTCG 402
DB 264 CATCCCAATGTGGAGCCAGTATTCTGGGGGAACATCAGATGACGCTCCAGAACAGTCG 323
QY 403 CTGTCATGGATCTGCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCATCCCAAGTGT 462
DB 324 CTGTCATGGATCTGCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCATCCCAAGTGT 383
QY 463 TAATCTGTAGTCGTGAGATGAACCTTGTGAAGTTACTGTCTACCTCAAGCTATTATTT 582
DB 384 TAATCTGTAGTCGTGAGATGAACCTTGTGAAGTTACTGTCTACCTCAAGCTATTATTT 503
QY 523 CCGGCTCCCGGATATTTCTGGGAGCTGCGTCTCTGCTGAGCCATCAAGCTATTATTT 582
DB 444 CCGGCTCCCGGATATTTCTGGGAGCTGCGTCTCTGCTGAGCCATCAAGCTATTATTT 503
QY 583 TGTTCGGAGCCAGCGACCTTCAAAGTCAGTGAGCATCTGGCTCTGACCCCAACAGAG 642
DB 504 TGTTCGGAGCCAGCGACCTTCAAAGTCAGTGAGCATCTGGCTCTGACCCCAACAGAG 563
QY 643 CAATGGGACTTTGACTTGGTGGCTACTCGGAGAGCTGAGCCCGCAAGCTGCAAC 702
DB 564 CAATGGGACTTTGACTTGGTGGCTACTCGGAGAGCTGAGCCCGCAAGCTGCAAC 623
QY 703 TGTAATATCTCACTGTGATTCGGTGTCTCCCAAGACACTGGAGGTTGGTATTATTTCCAGG 762
DB 624 TGTAATATCTCACTGTGATTCGGTGTCTCCCAAGACACTGGAGGTTGGTATTATTTCCAGG 683
QY 763 TGTAATATCAAGTTTACCGAGTTTAGGTTTTCATTCGCTACTTGGGGGCAAGTTGGACT 822

Db 684 TGTATTATCAAGTTTACCGAGTTTAGGTTTTCATTCGCTACTTGGGGCAAGTTGGACT 743
QY 823 TGGACTAGCAGCACCAGTCTTCTGAGCGCGACGTGTACTCTTACAATACGCTGTGCTG 882
Db 744 TGGACTAGCAGCACCAGTCTTCTGAGCGCGACGTGTACTCTTACAATACGCTGTGCTG 803
QY 883 CTGCGCGCGCTGCTGTGTGTGGCTGCAACTGCTGCGCGCTGTGTGTCTGCTGTAGAAG 942
Db 804 CTGCGCGCGCTGCTGTGTGTGGCTGCAACTGCTGCGCGCTGTGTGTCTGCTGTAGAAG 863
QY 943 AAAAAGAGGA 952
Db 864 AAAAAGAGGA 873
RESULT 7
AF121782 142742 bp DNA linear PRI 02-FEB-1999
LOCUS Homo sapiens chromosome 21q22.3 PAC 206A10, complete sequence.
DEFINITION
ACCESSION AF121782
VERSION AF121782.1 GI:4210991
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 142742)
AUTHORS Taudien, S., Dagand, E., Hildmann, T., Nordsiek, G., Drescher, B.,
Schattke, R., Weber, J., Schilling, M., Menzel, U., Yaspo, M.L. and
Rosenthal, A.
TITLE Direct Submission
JOURNAL Submitted (21-JAN-1999) Genome Analysis, Institute for Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES
source
1..142742
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="PAC 206A10"
3..1118
/evidence=not_experimental
repeat_region
/rpt_family="MIR"
repeat_region
/evidence=not_experimental
/rpt_family="MIR"
repeat_region
/evidence=not_experimental
/rpt_family="MIR"
exon
/evidence=not_experimental
/rpt_family="MIR"
/notes="GRAIL, score = 86.000%, comment = excellent shadow"
repeat_region
/evidence=not_experimental
/rpt_family="MIR"
repeat_region
/evidence=not_experimental
/rpt_family="MIR"
exon
/evidence=not_experimental
/rpt_family="MIR"
exon
/evidence=not_experimental
/rpt_family="MIR"
repeat_region
/evidence=not_experimental
/rpt_family="MIR"
repeat_region
/evidence=not_experimental
/rpt_family="MIR"
exon
/evidence=not_experimental
/rpt_family="MIR"
exon
/evidence=not_experimental
/rpt_family="MIR"
repeat_region
/evidence=not_experimental
/rpt_family="MIR"
repeat_region
/evidence=not_experimental
/rpt_family="MIR"
exon
/evidence=not_experimental
/rpt_family="MIR"


```
Query Match      24.4%; Score 302; DB 9; Length 142742;
Best Local Similarity 100.0%; Pred. No. 6.4e-168;
Matches 302; Conservative 0; Mismatches 0; Indels 0;
```

433	AGTTATGGGAGAGCTGTGCAATCCCAAGTGTTAACTCTTGTAAGTCGCTCGAGAATGAACCTTG	492
y		
4654	AGTTATGGGAGAGCTGTGTCATCCCAAGTGTTAACTCTTGTAAGTCGCTCGAGAATGAACCTTG	49713
b		
493	TGAAGTTACTTTGTCTACCTCACACTGAGCCCGGCTCCCGGATATTTCTGGGAGCTCG	552
y		
49714	TGAAGTTACTTTGTCTACCTCACACTGAGCCCGGCTCCCGGATATTTCTGGGAGCTCG	49773
b		
553	TCCTCTGGTCAGACCAATTCAGACTATATTTTGTTCGGAGCCACGACCTTCAAAGTGC	612
y		
49774	TCCTCTGGTCAGACCAATTCAGACTATATTTTGTTCGGAGCCACGACCTTCAAAGTGC	49833
b		
613	AGTGAGCATCTCGGCTCTGACCCACAGAGCAATGGGACTTTTGACTTGGCTGAGTACCTG	672
y		
49834	AGTGAGCATCTCGGCTCTGACCCACAGAGCAATGGGACTTTTGACTTGGCTGAGTACCTG	49893
b		
673	GAAGAGCCTGAAGCCCGCAAGTCTGCAACTGTAAATCTCACTGTGATTCGGTGTGCCCA	732
y		
49894	GAAGAGCCTGAAGCCCGCAAGTCTGCAACTGTAAATCTCACTGTGATTCGGTGTGCCCA	49953
b		
733	AG 734	
y		
49954	AG 49955	
b		

RESULT 8				
JOCUS	HS21C080	340000 bp	DNA	linear
DEFINITION	Homo sapiens chromosome 21 segment HS21C080.			
ACCESSION	AL163280 AF001735 BA000005			
VERSION	AL163280.2 GI:7717369			
KEYWORDS	.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata;			
	Mammalia; Eutheria; Primates;			
	Catarrhini; Hominoidea; Homo.			
	PRI 24-MAY-2000			

REFERENCE
AUTHORS

1 (Bases 1 to 340000)

Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T., Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E., Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K., Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R., Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A., Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K., Zaakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuama,S., Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordstiek,G., Hornischer,K., Brandt,P., Schaffe,M., Schoen,O., Desario,A., Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S., Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K., Gardner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and Yaspo,M.L.

TITLE
JOURNAL
Direct Submission
Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing Consortium * RIKEN Genomic Sciences Center, Human Genome Research Group * Institute of Molecular Biotechnology, Genome Analysis * Keio University School of Medicine, Dept. of Molecular Biology * GSF, Dept. of Genome Analysis * Max-Planck Institute for Molecular Genetics (addresses see below)
Responsible Person: [redacted]

* RIKEN Genomic Sciences Center, Human Genome Research Group, *
 * e-mail: sakaki@gs.riken.go.jp
 * URL: <http://hqp.gsc.riken.go.jp/>

and
* Institute of Molecular Biotechnology, Genome Analysis, *
Beutenbergstrasse 11, D-07745 Jena, Germany,
* e-mail: gscj-submit@genome.imb-jena.de
* URL: <http://genome.imb-jena.de/>
and
* Zeo Universität School of Medicine, Dept. of Molecular Biology, *

```

/note="(TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3005..3074
/note="(TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3473..19157
/gene="SH3BGR"
join(<3473..3565,12649..12678,15331..15460,18739..19157)
/gene="SH3BGR"
join(3473..1565,12649..12678,15331..15426)
/gene="SH3BGR"
/note="Accession No. X93498"
/codon_start=1
/product="21-Glutamic Acid Rich protein 21-GARP"
/protein_id="CA930445.1"
/db_xref="GI:7717370"
/db_xref="GOA:P55822"
/db_xref="SWISS-PROT:P55822"
/translators="GSEKAEGETEAKGSESDVGNLPEAQEKNEEGETATEETEE
IMAGEAGEAEETAEGETEAKGSESDVGNLPEAQEKNEEGETATEETEE
3473..3565
/gene="SH3BGR"
/number=4
5542..5829
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(6887..7067)
/note="L1MD3"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
complement(8482..8762)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(8865..8964)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
9131..9424
/note="AluSq"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(9723..9882)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(9989..10077)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(10078..10429)
/note="THE1B"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
complement(10430..10651)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
10980..11000
/note="(TTTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(11002..11282)
/note="AluSp"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(11315..11556)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED

```

```

repeat_region 11702..11753
/note="MIR"
/rpt_family="SINE/MIR"
/rpt_type=DISPERSED
12649..12678
/gene="SH3BGR"
/number=5
repeat_region complement(13115..13425)
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
repeat_region complement(13598..13892)

Query Match 24.4% Score 302; DB 9; Length 340000;
Best Local Similarity 100.0%; Pred. No. 6.9e-168; Indels 0; Gaps 0;
Matches 302; Conservative 0; Mismatches 0;

Qy 433 AGTTATGGGAGAGCTGTTCAATCCAGTGTAACTTTGTAGTCGCTGAGAAATGAACCTTG 492
Db 273834 AGTTATGGGAGAGCTGTTCAATCCAGTGTAACTTTGTAGTCGCTGAGAAATGAACCTTG 273893

Qy 493 TGAAGTTACTTGTCTACCCCTCACACTGGACCGGCTCCGGATATTTCTCTGGGAGCTCGG 552
Db 273894 TGAAGTTACTTGTCTACCCCTCACACTGGACCGGCTCCGGATATTTCTCTGGGAGCTCGG 273953

Qy 553 TCTCTGCTCAGCCATTCAAGCTATTATTTTGTTCGGAGCCGACGACCTTCAAGTGC 612
Db 273954 TCTCTGCTCAGCCATTCAAGCTATTATTTTGTTCGGAGCCGACGACCTTCAAGTGC 274013

Qy 613 AGTGAGCATCTGGCTCTGACCCACAGAGCAATGGGACTTTGACTTGCCTGGCTACCTG 672
Db 274014 AGTGAGCATCTGGCTCTGACCCACAGAGCAATGGGACTTTGACTTGCCTGGCTACCTG 274073

Qy 673 GAAGAGCTCAAGCGCGGCAAGTCTGCAACTGTAATCTCACTGTGATTCGGTGTCCCCA 732
Db 274074 GAAGAGCTCAAGCGCGGCAAGTCTGCAACTGTAATCTCACTGTGATTCGGTGTCCCCA 274133

Qy 733 AG 734
Db 274134 AG 274135

RESULT 9
CH179K04 182532 bp DNA linear PRI 16-OCT-2003
LOCUS Pan troglodytes chromosome 22 BAC CH251-179K04, complete sequence.
ACCESSION AL954228
VERSION AL954228.1 GI:37619870
KEYWORDS HTG.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 182532)

REFERENCE
AUTHORS The Chimpanzee Chromosome 22 Sequencing Consortium
CONSTRM Chimpanzee chromosome 22 genomic sequence
TITLE Unpublished
JOURNAL 2 (bases 1 to 182532)
REFERENCE 2 (bases 1 to 182532)
AUTHORS Schafhe, M., Berg, C., Conrad, A., Hornischer, K., Loehnert, T.H.,
Ludwig, M., Thies, S., Weber, K. and Bloecker, H.
TITLE Direct Submission
JOURNAL Submitted (03-JUN-2003) GSF, Dept. of Genome Analysis, Mascheroder
Weg 1, D-38124 Braunschweig, Germany, E-mail: info.genome@gsf.de
COMMENT The Chimpanzee Chromosome 22 Sequencing Consortium consists of :
Shanghai, China
Taiwan;
----- Genome Center
Center: GSF, Braunschweig
Center code: GSF
Web site: http://genome.gsf.de/
Contact: info.genome@gsf.de
----- Project Information

```

Direct Submission

Submitted (15-MAY-2003) Naruya Saitou, National Institute of Genetics (NIG), Division of Population Genetics; 1111 Yata, Mishima, Shizuoka 411-8540, Japan (E-mail: nsaitou@genes.nig.ac.jp, URL: <http://sayer.lab.nig.ac.jp/>, Tel: 81-55-981-6790, Fax: 81-55-981-6789)

The Chimpanzee Chromosome 22 Sequencing Consortium consists of:

- *Chinese National Human Genome Center at Shanghai, Shanghai, China;
- *GSF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute of Molecular Biotechnology, Jena, Germany; *KRIBB Genome Research Center, Daejeon, Korea;
- *Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
- *National Institute of Genetics, Mishima, Japan;
- *National Yang Ming University Genome Research Center, Taipei, Taiwan;
- *RIKEN Genomic Sciences Center, Yokohama, Japan.

----- Genomic Center

Center: National Institute of Genetics

Center code: NIG

Web site: <http://sayer.lab.nig.ac.jp/>

Contact: nsaitou@genes.nig.ac.jp

----- Project Information

Center project name: The Chimpanzee Chromosome 22 Sequencing Project

Center clone name: PTB-060F12

----- Summary Statistics

Sequencing vector: pUC118; 100% of reads

Chemistry: Dye-terminator Big Dye and dGMP; 100% of reads Assembly

Program: Phrap; version 0.990329

Consensus quality: 156,231 bases at least Q40

Consensus quality: 156,288 bases at least Q30

Consensus quality: 156,288 bases at least Q20

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30);

an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at one plasmid

subclone or more than one M13 subclone;

Source information:

The PTB1 chimpanzee BAC library was prepared from DNA isolated from cultured cells established from the blood of a single male chimpanzee.

Clones may be obtained from Asao Fujiyama and co-workers (<http://www.gsc.riken.go.jp>).

VECTOR: pKS145

Sequence Quality Assessment:

This entry has been annotated with sequence estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.

Neighboring clones: CH251-179K04 (left) and PTB-103H04 (right).

Location/Qualifiers

1. .156288

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/chromosome="22"

/clone="PTB-060F12"

/clone_lib="PTB1 chimpanzee BAC"

Ch 11.7%; Score 145; DB 9; Length 156288;

11 Similarity 100.0%; Pred. No. 5.6e-74;

1445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

145 GAGCTGGTCTCCTGCTAGCATTCAAGCTATTTATTTGTTCCGAGCCACGACCTT 604

QY 545 GAGCTGGTCTCCTGGTCAGCCATTCAAAGCTATTATTTTGTTCGGAGCCCCAGCGACCTT 604

```

Db      1 GAGTCGGTCTCCGGTCAGCCATTCAGCTATTATTGTTCCGGAGCCAGCGACCTT 60
Qy      605 CAAAGTCAGTGAGCATCTCGCTCTGACCCACAGACCAATGGACTTTCGACTTGGGTG 664
         |||||
Db      61 CAAAGTCAGTGAGCATCTCGCTCTGACCCACAGACCAATGGACTTTCGACTTGGGTG 120
         |||||
Qy      665 GCTACCTGGAGAGCCTGAAGGCC 689
         |||||
Db      121 GCTACCTGGAGAGCCTGAAGGCC 145

RESULT 11
AF045450 40205 bp DNA linear PRI 20-MAR-1998
LOCUS Homo sapiens chromosome 21q22.3 cosmid Q11M15, complete sequence.
ACCESSION AF045450
VERSION AF045450.1 GI:2895783
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 40205)
AUTHORS Taudien,S. and Rosenthal,A.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 40205)
AUTHORS Taudien,S., Nordsiek,G., Dagand,E., Hildmann,T., Drescher,B.,
Weber,J., Rosenthal,A. and Yaspo,M.L.
TITLE Submitted (29-JAN-1998) Genome Analysis, Institut for Molecular
JOURNAL Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES
     source
       1. 40205
           /organism="Homo sapiens"
           /mol_type="genomic DNA"
           /db_xref="taxon:9606"
           /chromosome="21"
           /map="21q22.3"
           /clone="cosmid Q11M15"
           /complement(293..586)
           /evidence=not_experimental
           /rpt_family="AluSc"
           /complement(643..964)
           /evidence=not_experimental
           /rpt_family="HERVL"
           /complement(1053..1176)
           /note="GenScan, score = 5.19%, comment = Internal_exon 124
bp frame: 2 phase: 1"
           /evidence=not_experimental
           /complement(1251..1859)
           /evidence=not_experimental
           /rpt_family="HERVL"
           /complement(2296..3220)
           /evidence=not_experimental
           /rpt_family="HERV16"
           /complement(2818..2889)
           /note="Xpound exon prediction, score = 82% (0%)"
           /evidence=not_experimental
           /complement(3221..3658)
           /evidence=not_experimental
           /rpt_family="MLT1C"
           /complement(3359..3395)
           /note="Xpound exon prediction, score = 84% (0%)"
           /evidence=not_experimental
           /complement(3874..4176)
           /evidence=not_experimental
           /rpt_family="AluJb"
           /complement(4330..4433)
           /evidence=not_experimental
           /rpt_family="L2"
           /complement(4585..4638)
           /note="GRAIL, score = 95.000%, comment = excellent shadow"
           /evidence=not_experimental

```

```

repeat_region complement(5113..7899)
                /evidence=not_experimental
                /rpt_family="L1PA14"
exon            complement(6119..6598)
                /note="GRAIL, score = 63.000%, comment = good"
                /evidence=not_experimental
                /complement(7927..8045)
                /evidence=not_experimental
                /rpt_family="FLAN_A"
repeat_region  8048..8178
                /evidence=not_experimental
                /rpt_family="L2"
exon            8771..8876
                /note="MZEF, score = 67.4%"
                /evidence=not_experimental
                /complement(8809..9187)
                /evidence=not_experimental
                /rpt_family="MER1A"
repeat_region  8837..8876
                /note="GRAIL, score = 55.000%, comment = good shadow"
                /evidence=not_experimental
                /complement(9113..9213)
                /note="GRAIL, score = 55.000%, comment = good"
                /evidence=not_experimental
                /complement(9190..9258)
                /evidence=not_experimental
                /rpt_family="L1MC3"
repeat_region  9259..9647
                /evidence=not_experimental
                /rpt_family="L1MC3"
repeat_region  complement(10138..10427)
                /evidence=not_experimental
                /rpt_family="MLT2B"
repeat_region  10862..10875
                /evidence=not_experimental
                /rpt_family="MER58A"
exon            complement(11072..11156)
                /note="GRAIL, score = 86.000%, comment = excellent"
                /evidence=not_experimental
                /complement(11072..11129)
                /note="Xpound exon prediction, score = 89% (0%)"
                /evidence=not_experimental
                /complement(11478..11597)
                /evidence=not_experimental
                /rpt_family="L2"
exon            11570..11704
                /note="MZEF, score = 66.1%"
                /evidence=not_experimental
                /complement(11570..11704)
                /note="GRAIL, score = 99.000%, comment = excellent shadow"
                /evidence=not_experimental
                /complement(11687..11735)
                /evidence=not_experimental
                /rpt_family="MIR"
repeat_region  complement(11774..12320)
                /evidence=not_experimental
                /rpt_family="MLT1E"
exon            complement(11828..11949)
                /note="GRAIL, score = 50.000%, comment = good"
                /evidence=not_experimental
                /complement(12623..12843)
                /evidence=not_experimental
                /rpt_family="AluSc"
repeat_region  13102..13457
                /evidence=not_experimental
                /rpt_family="THE1B"
exon            complement(13481..13519)
                /note="GRAIL, score = 65.000%, comment = good"
                /evidence=not_experimental
                /complement(15065..15188)
                /evidence=not_experimental
                /rpt_family="HERVL"
repeat_region  15130..15378
                /evidence=not_experimental
                /rpt_family="HERVL"

```

/note="GC score = 10.60 (249bp); Region: GC content"
/evidence=not_experimental
15369..16268
/rpt_family="MERVL"
complement(16623..16714)
/note="GRAIL, score = 72.000%, comment = good"
/evidence=not_experimental
16661..16730
/note="WZEF, score = 93.5%"
/evidence=not_experimental
16880..17318
/evidence=not_experimental
/rpt_family="MLT1C"
17326..17436
/evidence=not_experimental
/rpt_family="LTR16C"
complement(17540..17675)
/evidence=not_experimental
/rpt_family="MIR"
complement(17805..17864)
/evidence=not_experimental
/rpt_family="MLT1C"
18076..18098
/note="GRAIL, score = 45.000%, comment = marginal shadow"
complement(18219..18349)
/note="GenScan, score = 2.99%, comment = Internal_exon 131
bp frame: 1 phase: 2"
/evidence=not_experimental
complement(18717..19211)
/evidence=not_experimental
/rpt_family="MLT1C"
complement(19252..19399)
/note="GRAIL, score = 71.000%, comment = good"
/evidence=not_experimental
19269..19350
/note="GRAIL, score = 81.000%, comment = excellent shadow"
complement(19379..19399)
/note="Xpound exon prediction, score = 60% (0%)
/evidence=not_experimental
complement(19501..19881)
/note="WZEF, score = 50.6%"
/evidence=not_experimental
19945..20020
/note="homology = 100.00%, score = 38, counts = 2"
/evidence=not_experimental
/rpt_type=tandem
/rpt_unit="ttctgcacactctctgttgagaggtcggtttatcgc"
20223..20643
/evidence=not_experimental
/rpt_family="MLT1E"
20645..20910
/evidence=not_experimental
/rpt_family="MLT1F"
20961..21007
/evidence=not_experimental
/rpt_family="MLT1F"
21483..21781
/evidence=not_experimental
/rpt_family="AluCb"
complement(21504..21756)
/note="WZEF, score = 79.4%"
/evidence=not_experimental
complement(22325..22806)
/evidence=not_experimental

Query Match 9.6%; Score 119; DB 9; Length 40205;
Best Local Similarity 100.0%; Pred. No. 1.8e-58;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTGTGAGTCCAGCCAAACAGTGTGGATTCAGTTTCTTAGGTGCCATACAAAGCACCAT 60

Db 18694 AGGTGTGAGTCCAGCCAAACAGTGTGGATTCAGTTTCTTAGGTGCCATACAAAGCACCAT 18753
Qy 61 AACCTGGTGGCTTAGAACAAATGAAAGGCAATTGCTCAGGTTCCAGAAAGCTGTAGGTT 119
Db 18754 AACCTGGTGGCTTAGAACAAATGAAAGGCAATTGCTCAGGTTCCAGAAAGCTGTAGGTT 18812
RESULT 12
AF064860 Homo sapiens chromosome 21 clone PAC 70124 map 21q22.3, complete
LOCUS AF064860 170121 bp DNA linear PRI 05-MAR-2002
DEFINITION sequence.
ACCESSION AF064860
VERSION AF064860.2 GI:18958624
KEYWORDS HTG; HTGS DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 170121)
AUTHORS Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
Park,H.-S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.-K., Soeda,E.,
Onki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
Patterson,D., Reichwald,K., Rump,A., Schillhabel,M.B., Schudy,A.,
Zimmermann,W., Rosenthal,A., Kudoh,J., Kawasaki,K., Asakawa,S.,
Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S.,
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,
Horisicher,K., Brandt,P., Scharfe,M., Schoen,O., Desario,A.,
Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S.,
Hennig,S., Riesselmann,L., Dagand,E., Haaf,T., Wehrmeyer,S.,
Porzym,K., Gardiner,K., Nizetic,D., Francis,F., Lehrach,H.,
Reinhardt,R. and Yaspo,M.Laure.
The DNA sequence of human chromosome 21
Nature 405 (6784), 311-319 (2000)
2089799
10830953
REFERENCE 2 (bases 1 to 170121)
AUTHORS Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.
Direct Submission
TITLE Submitted (12-MAY-1998) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
JOURNAL
REFERENCE 3 (bases 1 to 170121)
AUTHORS Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L., Rosenthal,A., Yaspo,M.-L. and
Rosenthal,A.
Direct Submission
TITLE Submitted (27-FEB-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
JOURNAL
REFERENCE 4 (bases 1 to 170121)
AUTHORS Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.
Direct Submission
TITLE Submitted (05-MAR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
JOURNAL
COMMENT On Feb 27, 2002 this sequence version replaced gi:3171153.
FEATURES
Location/Qualifiers
source
1..170121
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="PAC 70124"

ORIGIN

Query Match 9.6%; Score 119; DB 9; Length 170121;
Best Local Similarity 100.0%; Pred. No. 2e-58;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTGTGAGTCCAGCCAAACAGTGTGGATTCAGTTTCTTAGGTGCCATACAAAGCACCAT 60

```

|||||
Db 159444 AGGTGTGAGTCCAGCCACAGTGGATCAGTTTCTTAGGCTGCATACCAAGCACCAT 159503
QV 61 AACCTGGTGGCTTAGACAATGGAAGGCACTTGTCTACGGTTCAGAGCTGTAGGTT 119
|||||
Db 159504 AACCTGGTGGCTTAGACAATGGAAGGCACTTGTCTACGGTTCAGAGCTGTAGGTT 159562

RESULT 13
AF064857 199665 bp DNA linear PRI 23-JUN-1998
LOCUS Homo sapiens BAC derived from chromosome 21q22.3, complete
DEFINITION sequence, containing PEP19 (PCP4) gene.
ACCESSION AF064857
VERSION AF064857.1 GI:3171149
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Taudien,S., Nordsiek,G., Korenberg,J., Drescher,B., Weber,J.,
Schattevoy,R. and Rosenthal,A.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-1998) Genome Analysis, Institute for Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES
source
1..199665
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
complement(1..383)
repeat_region /evidence=not experimental
/rpt_family="L1"
747..814
repeat_region /evidence=not experimental
/rpt_family="MIR"
complement(1345..1500)
repeat_region /evidence=not experimental
/rpt_family="L1PA2"
1499..1588
repeat_region /evidence=not experimental
/rpt_family="L1PA2"
complement(1584..2179)
repeat_region /evidence=not experimental
/rpt_family="L1PA2"
2174..3405
repeat_region /evidence=not experimental
/rpt_family="L1PA2"
3388..3454
repeat_region /evidence=not experimental
/rpt_family="L1p"
3968..4086
repeat_region /evidence=not experimental
/rpt_family="MIR"
complement(6245..6668)
repeat_region /evidence=not experimental
/rpt_family="L1ME3A"
6671..6969
repeat_region /evidence=not experimental
/rpt_family="AluJb"
complement(7057..7376)
repeat_region /evidence=not experimental
/rpt_family="MER2"
complement(7387..7491)
repeat_region /evidence=not experimental
/rpt_family="L1ME3A"
7781..8020
repeat_region /evidence=not experimental
/rpt_family="MER58A"
8650..8823
repeat_region /evidence=not experimental
/rpt_family="THE1B"

/evidence=not experimental
/rpt_family="L1PA5"
complement(10456..10720)
/evidence=not experimental
/rpt_family="L2"
complement(13514..13814)
/evidence=not experimental
/rpt_family="AluSx"
complement(13823..14116)
/evidence=not experimental
/rpt_family="AluJo"
complement(15662..15785)
/evidence=not experimental
/rpt_family="MER5A"
complement(18994..19347)
/evidence=not experimental
/rpt_family="THE1B"
19378..19470
/notes="homology = 81.70%, score = 28, counts = 3"
/evidence=not experimental
/rpt_type=tandem
/rpt_unit="agtatctattctgattgggcagtgctcatcac"
complement(20371..20556)
/evidence=not experimental
/rpt_family="MER3"
21213..21500
/evidence=not experimental
/rpt_family="AluSx"
22291..22567
/evidence=not experimental
/rpt_family="MER39b"
complement(22583..23201)
/evidence=not experimental
/rpt_family="L1PA2"
23202..23471
/evidence=not experimental
/rpt_family="MER39b"
24223..24454
/evidence=not experimental
/rpt_family="MSTB-internal"
24455..24889
/evidence=not experimental
/rpt_family="MSTB"
complement(25434..25767)
/evidence=not experimental
/rpt_family="MER58B"
25993..26294
/evidence=not experimental
/rpt_family="AluSq"
27222..27283
/evidence=not experimental
/rpt_family="MIR"
complement(28932..29335)
/evidence=not experimental
/rpt_family="MSTA"
complement(29462..30001)
/evidence=not experimental
/rpt_family="MER34"
complement(30308..30676)
/evidence=not experimental
/rpt_family="MLT1B"
31012..31364
/evidence=not experimental
/rpt_family="MLT1A2"
complement(32507..32738)
/evidence=not experimental
/rpt_family="MSTD"
complement(32766..32980)
/evidence=not experimental
/rpt_family="MSTD"
33468..33678
/evidence=not experimental
/rpt_family="THE1B"

```



```

/note="assembly_fragment-clone_end:T7-vector side:left
assembly_fragment-clone_end:SP6-vector_side:right"

ORIGIN
Query Match          5.8%; Score 72; DB 9; Length 192219;
Best Local Similarity 100.0%; Pred. No. 2.7e-30;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 GGGAACTTCACCTCGGAGATGATCATCCACATGTGGAGCCAGTGATTCCGGGAACATC 379
      |||||
Db 191622 GGGAACTTCACCTCGGAGATGATCATCCACATGTGGAGCCAGTGATTCCGGGAACATC 191681

QY 380 AGATGCAGCCTC 391
      |||||
Db 191682 AGATGCAGCCTC 191693

RESULT 15
AB090820/c
LOCUS      AB090820          6470 bp      DNA      linear      INV 25-MAR-2003
DEFINITION Anopheles gambiae retrotransposon R7Ag1 DNA, complete sequence.
ACCESSION  AB090820
VERSION     AB090820.1 GI:28569877
KEYWORDS
SOURCE      Anopheles gambiae (African malaria mosquito)
ORGANISM    Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
            Anopheles.
REFERENCE   1
AUTHORS    Kojima,K.K. and Fujiwara,H.
TITLE      Evolution of Target Specificity in R1 Clade Non-LTR
            Retrotransposons
JOURNAL    Mol. Biol. Evol. 20 (3), 351-361 (2003)
MEDLINE    22531580
PUBMED    12644555
REFERENCE   2 (bases 1 to 6470)
AUTHORS    Kojima,K.K. and Fujiwara,H.
TITLE      Direct Submision
JOURNAL    Submitted (26-AUG-2002) Kenji K Kojima, University of Tokyo,
            Department of Integrated Biosciences, Graduate School of Frontier
            Sciences, Bioscience Building 501, 5-1-5 Kashiwanoha, Kashiwa,
            Chiba 277-8562, Japan (E-mail:kk27513@mail.ecc.u-tokyo.ac.jp,
            Tel:81-4-7136-3661, Fax:81-4-7136-3660)
FEATURES
            Location/Qualifiers
            1..6470
               /organism="Anopheles gambiae"
               /mol_type="genomic DNA"
               /db_xref="taxon:7165"
            repeat_region
               1..6470
               /note="non-LTR retrotransposon"
               /transposon="R7Ag1"
               959..2542
               /note="ORF1"
               /codon_start=1
               /product="gag-like protein"
               /protein_id="BAC57915.1"
               /db_xref="GI:28569878"
               /translation="MDKQLRGRTISVDERPAVIRKLGSEKKLGTIVESPSSAGVPAR
               TWATGVKSAGTATKLTATPSTGVSRVRLADAKADNETTVGIVKRLBEQQLRLQ
               MEANSQLKEAREAREADARAEHRELRKEKELFNALLAQTLTGTSGARLES
               QQLERQELLRRMESQEQQLQDQQRQQRQQRQQRQQLPAQWPTVOQS
               VRAQRQVTESSASSVDEAGTWVEVVRGNKONGVNLPOOSACROPAPRHOOQ
               PQHQVQQQQRGHIHQEKKRRPRKRDPDEIVVPAGVSEKEMVYKIRNPIADPQ
               ROIGVGRTPRHLLPLSRDVSALKDIIQEVIGERSVTVTEMAEVLTGIDNM
               IDEEAIKKALMTLQSLVATVNLWERDMTKRVRUFRABEVLKDRRLLEGTY
               CSVHEAPKVSGLQTRFCRLGRGHIAATCTGDRSKRCLRCGQDQTHKSGCTNEVKCM
               LCGAHRIGAAACGGQPSN"
               2544..6212
               /note="ORF2"
               /codon_start=1
               /product="reverse transcriptase"
               /protein_id="BAC57916.1"

```

```

/db_xref="GI:28569879"
/translation="MEVLIQNVNRSRQAQDLAINTMVRADVCLMVELHSHVPRNNGN
WVADRDGKVAIIASSEYFVQQVSVTQSGIAARINDVLFICCYVSPSAGVSEFEV
MQRIDVLARGHPRVVFAGDLNMAHMTAGSCKRTNAKGAHVQLVDSLGLVINTGTAPT
FLGNVAPSVVDVAFASSIAGVNTIPEHRWRIISYYSNHYVIRFVAGVLLQRPAA
ADSRQGPSPTRSGTRWRTHFDAELFGVALDVASPTERTVTSAESLERVMTLEQDAA
MARVFPQSGHGRPAYVWTPAIVELCENCLAKERLEAAIDEEFQIAASDLLOVRTA
LDSITTSKKEHFEILLGLAEDETGWYRNVLRSLSGSWTARERDSSVLEIVSTLF
POHPVDVPASPOQVLERGEBEPVRDVNEQELLDIASLNERKAPGLDGVNAAITAA
IRKHTDIFKKLFOECLDNERFPDEWKQKLLALPKPKQPGPLGASSFRPILLLNPGKV
YERLLSRINDVIEDPESPRLAENQYFRGRSTVQAIQLVVDAGSHAMSFGRTNRRD
KRCLLVVALDVRAFNTASWCQIATALEDKGVRQLNRLRDYFANRELVTYTDAGPV
TRVTAGVPOGSIILGPTLWNIMYDGLVRVELPEGASVIGVADDIVMARGCTPOEAL
VARQAVDAIAAWMEDHHLQAPKTEGVMISLLRGGQLKVPFRVGDTHIHSKQIRYL
GVQIHDLHLSWKPHVELSTAKALRVGVVTAVMRNHSQVQAKRLLAAVESIIIRYAA
PWSSEATDLOWCQKQKLAQVORPLARGVTSSVAVYETGVALAGLVFRLVREDARC
HRRLLAAPGASRKDIRLEERQGTQFQEWQRAWDAATAAPTASRVASVAHVMIPDLHWM
SRHGEVDPHLSQVLTGHYFREYLVHVCFGAPGAECPCPGSVESVAHVLFQCEVPHE
IRVELLYGTSDPVNNLGMKLLSPERNWSIQEARKITKVLQQLWREDELQNLQ
AHLAALPTRAAADVADGPLDGEQVSDGVAELFRSNRGRARRTRRRRABERVEURLA
SMAAABEREREDSILMAAVRAEAGEAPPEIPMRRGLPSPPTVTRARHERRLYLQRL
YRQAREGLTPTVPHGRNRRRSAPSADTIRRRMRREMERLURTRARVPSNQGVRE
VLSADALAAITEATTSGR"

```

ORIGIN

Query Match 2.0%; Score 25; DB 3; Length 6470;
 Best Local Similarity 100.0%; Pred. No. 0.027;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 912 GCTGCTGCCGTGTGTGTTCTGCTG 936
 |||||
 Db 1574 GCTGCTGCCGTGTGTGTTCTGCTG 1550

Search completed: September 15, 2004, 03:47:27
 Job time : 5100.01 secs

This Page Blank (uspto)

Result No.	Score	Query Match	Length	DB	ID	Description
1	1240	100.0	1240	6	ABK13030	DNA encod
2	1125	90.7	2051	9	ADB62950	Human GDN
3	1074	86.6	1175	6	ABK13028	DNA encod
4	976	78.7	1168	6	ABK13029	DNA encod
5	739	59.6	1139	6	ABK13031	DNA encod
6	728	58.7	1392	5	AAS92356	DNA encod
7	268	21.6	474	8	ACH16130	DNA encod
8	144	11.6	401	4	AAI36582	Human adu
9	25	2.0	357	5	AAI36582	Probe #52
c 10	24	1.9	6507	4	ABL104335	Probe #43
c 11	24	1.9	59967	4	ABL15493	Drosophil
12	23	1.9	11195	6	ABK13032	DNA encod
13	22	1.8	141	2	AAV99359	5' PCR pr
14	22	1.8	767	6	ABN99162	Arabidops
15	22	1.8	1200	2	AAV35364	Human GDN
16	22	1.8	1200	2	AAV35365	Human GDN
17	22	1.8	1203	3	AAA12547	DNA encod
18	22	1.8	1699	2	AAV00251	DNA encod
19	22	1.8	1792	3	AAZ29104	Human GFR
20	22	1.8	1809	2	AAV99933	Human Ret
21	22	1.8	1829	3	AAZ93702	Glia1 cel
22	22	1.8	1829	3	AAA88519	PRO538 DN
23	22	1.8	1829	3	AAA77617	Human PRO

Key	Location/Qualifiers
CDS	104..1327
	/*tag= a
	/product= "Clone PLACES0177880 protein"

EP1308459-A2.

07-MAY-2003 .
28-MAR-2002: 2002EP-00007401

05-NOV-2001; 2001JP-00379298.
25-JAN-2002; 2002JIS-00250070

(HELI-) HELIX RES INST.
(REAS-) RES ASSOC BIOTECHNOLOGY.

Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y.

WPI; 2003-450961/43.
P-PSDB; ADB64920.

New polynucleotides and polypeptides, useful for developing a diagnostic marker or medicines for regulation of their expression and activity, or as targets of gene therapy.

Claim 1; Page: 222pp; English.

The invention discloses a polynucleotide comprising a sequence selected from 1970 fully defined nucleotide sequences which encode novel polypeptides. Also claimed is a polypeptide encoded by the polynucleotide or its partial peptide, an antibody binding to the polypeptide or peptide of the polynucleotide, immunologically assaying the polypeptide or peptide of the polynucleotide by contacting the polypeptide or peptide with the antibody of the encoded protein, and observing the binding between the two, a transformant carrying the polynucleotide in an expressible manner and an antisense polynucleotide. The oligonucleotide is useful as a primer for synthesising the polynucleotide, or as a probe for detecting the polynucleotide. The polynucleotides and encoded proteins are useful as pharmaceutical agents and many disease-related genes may be included in them, for developing a diagnostic marker or medicines for regulation of their expression and activity, or as targets of gene therapy. The genes are involved in tissue and/or cell regeneration. Membrane proteins, signal transduction-related proteins, transcription-related proteins, disease-related proteins and genes encoding them can be used as indicators for diseases (e.g. osteoporosis, neurological diseases, cancer, tumours. The cDNA may be used to regulate the activity or expression of the encoded protein to treat diseases. The sequence presented is a cDNA of the invention. Note: Some of the sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office.

Sequence 2051 BP; 580 A; 463 C; 474 G; 534 T; 0 U; 0 Other;

```

every Match      90.7%; Score 1125; DB 9; Length 2051;
1st Local Similarity 100.0%; Pred. No. 0;
Mismatches 1125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

116 GGTTCTGGGTCCTGGTAATGAAGTCATAGAAGGCCCCAGAAATGCAACAGTCCTGAAGGC 175
|||||
203 GGTTCTGGGTCCTGGTAATGAAGTCATAGAAGGCCCCAGAAATGCAACAGTCCTGAAGGC 262

176 TCCAGGCTCGCTTCAACTGCACCGTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTC 235

263 TCCAGGCTCGCTTCAACTGCACCGTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTC 322

236 AGTGCATGTTGGTCTAAGCGTCAGGGCCCATGGAGGCCCATCATCACCAATGACCGCTTC 295

296 ACCTCTCAGAGGTACGACCAGGGGGGAACCTTACCTCGGAGATGATCATCCACAATGTG 355

[illegible]

RESULT 3	
ABK13028	
ID	ABK13028 standard; cDNA; 1175 BP.
XX	
AC	ABK13028;
XX	
DT	23-APR-2002 (first entry)
XX	
DE	DNA encoding human B7-like protein

XX Human; B7-like protein; B7-L; antiinfertility; gynaecological;
KW antitumor; cytostatic; immunosuppressive; antiarthritic; antirheumatic;
KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;
KW antidiabetic; haemostatic; antichyroid; antitumor; antiallergic;
KW antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;
KW reproductive disorder; graft versus host disease; autoimmune disease;
KW toxic shock syndrome; allergy; nephropathy; skin disorder;
KW endocrinopathy; lymphoproliferative disorder; gene; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 27..1175
FT /tag= a
FT /product= "B7-like protein, B7-L_h1"
XX
FN WO200200710-A2.
PD 03-JAN-2002.
XX
XX 28-JUN-2001; 2001WO-US020719.
XX
XX 28-JUN-2000; 2000US-0214512P.
PR 28-NOV-2000; 2000US-00729264.
XX
XX (AMGE-) AMGEN INC.
XX
XX Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
XX WPI; 2002-130881/17.
DR P-PSDB; AAU75540.
XX
XX New B7-like polypeptides, polynucleotides and their modulators, useful
PT for diagnosing, preventing and treating reproductive, immune and
PT proliferative disorders, e.g. cancer and arteriosclerosis.
XX
XX Claim 1; Fig 1; 135pp; English.
XX
XX The invention relates to an isolated B7-like (B7-L) polypeptide (I). The
CC polypeptide, polynucleotide encoding it and antibody against (I) are
CC useful for treating B7-like polypeptide-related disease, disorders or
CC conditions including reproductive disorders (e.g. infertility,
CC miscarriage, preterm labour and delivery and endometriosis) and
CC proliferative disorders. Antibodies, soluble proteins comprising
CC extracellular domains and other regulators of B7-L polypeptides are
CC useful for enhancing the immune response to tumours. (I) plays a role in
CC growth and maintenance of cancer cells based on the observation of
CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
CC polypeptide. Hence modulators of (I) are useful for the treatment of
CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response in
CC allograft transplantation, graft versus host disease, T-cell dependent B-
CC cell mediated diseases and autoimmune diseases. B7-L molecules are useful
CC for alleviating the symptoms associated with diseases involving chronic
CC immune cell dysfunction or to treat autoimmune diseases such as systemic
CC lupus erythematosus, rheumatoid arthritis, multiple sclerosis, diabetes,
CC immune thrombocytopenic purpura and psoriasis, chronic inflammatory
CC disease such as inflammatory bowel disease (Crohn's disease and
CC ulcerative colitis), Grave's disease, Hashimoto's thyroiditis and
CC diabetes mellitus. They are also useful as immunosuppressive agents for
CC bone marrow and organ transplantation or to prolong graft survival. B7-L
CC molecules are also useful for diagnosis and treatment of diseases
CC involving abnormal cell proliferation, including arteriosclerosis and
CC vascular stenosis. Antagonists of B7-L polypeptides are useful for
CC alleviation of toxic shock syndrome or allosensitisation due to blood
CC transfusions, and for treatment of allergy, asthma and hypersensitivity
CC reactions, nephropathies (e.g. glomerulonephritis), skin disorders
CC (pemphigus and pemphigoid), endocrinopathies (Grave's disease), various
CC pneumopathies (extrinsic alveolitis), vasculopathies, coeliac disease,
CC anaemia, thrombocytopenias, Guillain-Barre syndrome and myasthenia
CC gravis, and lymphoproliferative disorders such as multiple myeloma. The

CC present sequence represents the coding sequence of human B7-L_h1
XX
SQ Sequence 1175 BP; 295 A; 307 C; 286 G; 287 T; 0 U; 0 Other;
Query Match 86.6%; Score 1074; DB 6; Length 1175;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1124; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 116 GGTCTCTGGTCTGGTAATGAAGTCATAGAGGCCCCAGAGTCAACAGTCTCTGAAGGGC 175
DB 51 GGTCTCTGGTCTGGTAATGAAGTCATAGAGGCCCCAGAGTCAACAGTCTCTGAAGGGC 110
QY 176 TCCAGGGCTCGCTTCAACTGCACCGCTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTC 235
DB 111 TCCAGGGCTCGCTTCAACTGCACCGCTCTCCAGGGCTGGAAGTCTCATCATGTGGGCTCTC 170
QY 236 AGTGACATGGTGGTCTGAAGCGTCAGGCCCATGAGGCCCATCATCAACATGACCGCTTC 295
DB 171 AGTGACATGGTGGTCTGAAGCGTCAGGCCCATGAGGCCCATCATCAACATGACCGCTTC 230
QY 296 ACTCTCTCAGAGGTACGACCCAGGGCGGGAATTTCACTCGGAGATGATCATCCACAATGTG 355
DB 231 ACTCTCTCAGAGGTACGACCCAGGGCGGGAATTTCACTCGGAGATGATCATCCACAATGTG 290
QY 356 GAGCCCAAGTATTCGGGGAAACATCAGATCAGCGCTCCAGAACAGTCCGCTGCAATGATCT 415
DB 291 GAGCCCAAGTATTCGGGGAAACATCAGATCAGCGCTCCAGAACAGTCCGCTGCAATGATCT 350
QY 416 GCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTTCATTCACAGTGTTAATCTTGTAGTC 475
DB 351 GCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTTCATTCACAGTGTTAATCTTGTAGTC 410
QY 476 GCTGAGAATGAACCTTGTGAAGTTACTTCTTACCTCACAAGTGGAGCCCGCTCCCGGAT 535
DB 411 GCTGAGAATGAACCTTGTGAAGTTACTTCTTACCTCACAAGTGGAGCCCGCTCCCGGAT 470
QY 536 ATTTCTCGGAGCTCGGTCTCTGTGTCAGCCCAATCAAGCTATTAATTTTTCGAGAGCCC 595
DB 471 ATTTCTCGGAGCTCGGTCTCTGTGTCAGCCCAATCAAGCTATTAATTTTTCGAGAGCCC 530
QY 596 ACGGACCTTCAAAGTGCAGTGCAGTCTGCGTCTGACCCCAAGAGCAATGGGACTTTG 655
DB 531 ACGGACCTTCAAAGTGCAGTGCAGTCTGCGTCTGACCCCAAGAGCAATGGGACTTTG 590
QY 656 ACTTGTGCTGCTACCTTGAAGAGCTTGAAGGCCGCAAGTCTGCAACTGTAAATCTCACT 715
DB 591 ACTTGTGCTGCTACCTTGAAGAGCTTGAAGGCCGCAAGTCTGCAACTGTAAATCTCACT 650
QY 716 GTGATTCGGTGTCCCCAAGACATGGAGTGGTATTAAATTAATTCAGGTGTATTATCAAGT 775
DB 651 GTGATTCGGTGTCCCCAAGACATGGAGTGGTATTAAATTAATTCAGGTGTATTATCAAGT 710
QY 776 TTACCGAGTTTAGGTTTTTTCATTTGCTGCTGCGGCAAGTTGGACTTGGACTAGCAGGC 835
DB 711 TTACCGAGTTTAGGTTTTTTCATTTGCTGCTGCGGCAAGTTGGACTTGGACTAGCAGGC 770
QY 836 ACCATGCTTCTCAGCCCGGACGTGACTCTTACAATACGCTGTGCTGCTGCGGCGCTCGT 895
DB 771 ACCATGCTTCTCAGCCCGGACGTGACTCTTACAATACGCTGTGCTGCTGCGGCGCTCGT 830
QY 896 TGTGTGGTGCACATGCTGCTGCGGTGTTGTTTCTGCTGTAGAGAAAAAGAGATTT 955
DB 831 TGTGTGGTGCACATGCTGCTGCGGTGTTGTTTCTGCTGTAGAGAAAAAGAGATTT 890
QY 956 CGTATTCAATTTCAAAGGAAATCTGAAGAAGAGACAGACAAAGAAATCTGACAGAA 1015
DB 891 CGTATTCAATTTCAAAGGAAATCTGAAGAAGAGACAGACAAAGAAATCTGACAGAA 950
QY 1016 AGTGAAATGAAATCTCCGGCTTACAAATTCAGATGAAACAAAGACACAGAAACCGCTTCT 1075
DB 951 AGTGAAATGAAATCTCCGGCTTACAAATTCAGATGAAACAAAGACACAGACACCGCTTCT 1010
QY 1076 CTCCTCCCAAAATCTGTGTAATCCAGTGTATCTGAAACAAAGAAACAGTAGCTGTGGCCCT 1135

Db 1011 CTCCTCCCAATCTCTGTGAATCCAGTGATCTCTGAACAAAGAAACAGTACTGTGGCCCT 1070
 QY 1136 CCTCACCAGCGGCTGATCAAGCTCCACCCAGGCCAGCAAGTATCCACAGGCTTCTTTT 1195
 Db 1071 CCTCACCAGCGGCTGATCAAGCTCCACCCAGGCCAGCAAGTATCCACAGGCTTCTTTT 1130
 QY 1196 AATCTGCCAGTCTCTGAGAAGGTCAGTAATACAACTGATATAG 1240
 Db 1131 AATCTGCCAGTCTCTGAGAAGGTCAGTAATACAACTGATATAG 1175

RESULT 4

ABK13029

ID ABK13029 standard; cDNA; 1168 BP.

XX AC ABK13029;

XX DT 23-APR-2002 (first entry)

XX XX DNA encoding human B7-like protein, B7-L_h2.

XX DE Human; B7-like protein; B7-L; antiinfertility; gynaecological;

XX KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;

XX KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;

XX KW antidiabetic; haemostatic; antithyroid; antiulcer; antiallergic;

XX KW antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;

XX KW reproductive disorder; graft versus host disease; autoimmune disease;

XX KW toxic shock syndrome; allergy; nephropathy; skin disorder;

XX KW endocrinopathy; lymphoproliferative disorder; gene; ss.

XX OS Homo sapiens.

XX XX

XX FH Key Location/Qualifiers

XX FT CDS 8..1168

XX FT /*tag= a

XX FT /product= "B7-like protein, B7-L_h2"

XX XX

XX PN WO200200710-A2.

XX XX

XX PD 03-JAN-2002.

XX PF 28-JUN-2001; 2001WO-US020719.

XX PR 28-JUN-2000; 2000US-0214512P.

XX PR 28-NOV-2000; 2000US-00729264.

XX XX

XX PA (AMGE-) AMGEN INC.

XX XX

XX PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;

XX XX

XX DR WPI; 2002-130881/17.

XX DR P-PSDB; AAU75541.

XX XX

XX PT New B7-like polypeptides, polynucleotides and their modulators, useful

XX PT for diagnosing, preventing and treating reproductive, immune and

XX PT proliferative disorders, e.g. cancer and arteriosclerosis.

XX XX

XX PS Claim 1; Fig 2; 135pp; English.

XX XX

XX CC The invention relates to an isolated B7-like (B7-L) polypeptide (I). The

XX CC polypeptide, polynucleotide encoding it and antibody against (I) are

XX CC useful for treating B7-like polypeptide-related disease, disorders or

XX CC conditions including reproductive disorders (e.g. infertility,

XX CC miscarriage, preterm labour and delivery and endometriosis) and

XX CC proliferative disorders. Antibodies, soluble proteins comprising

XX CC extracellular domains and other regulators of B7-L polypeptides are

XX CC useful for enhancing the immune response to tumours. (I) plays a role in

XX CC growth and maintenance of cancer cells based on the observation of

XX CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L

XX CC polypeptide. Hence modulators of (I) are useful for the treatment of

XX CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,

XX CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide

XX CC

CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response in
 CC allograft transplantation, graft versus host disease, T-cell dependent B-
 CC cell mediated diseases and autoimmune diseases. B7-L molecules are useful
 CC for alleviating the symptoms associated with diseases involving chronic
 CC immune cell dysfunction or to treat autoimmune diseases such as systemic
 CC lupus erythematosus, rheumatoid arthritis, multiple sclerosis, diabetes,
 CC immune thrombocytopenic purpura and psoriasis, chronic inflammatory
 CC disease such as inflammatory bowel disease (Crohn's disease and
 CC ulcerative colitis), Grave's disease, Hashimoto's thyroiditis and
 CC diabetes mellitus. They are also useful as immunosuppressive agents for
 CC bone marrow and organ transplantation or to prolong graft survival. B7-L
 CC molecules are also useful for diagnosis and treatment of diseases
 CC involving abnormal cell proliferation, including arteriosclerosis and
 CC vascular restenosis. Antagonists of B7-L polypeptides are useful for
 CC alleviation of toxic shock syndrome or allosensitisation due to blood
 CC transfusions, and for treatment of allergy, asthma and hypersensitivity
 CC reactions, nephropathies (e.g. glomerulonephritis), skin disorders
 CC (pemphigus and pemphigoid), endocrinopathies (Grave's disease), various
 CC pneumopathies (extrinsic alveolitis), vasculopathies, coeliac disease,
 CC anaemia, thrombocytopenias, Guillain-Barre syndrome and myasthenia
 CC gravis, and lymphoproliferative disorders such as multiple myeloma. The
 CC present sequence represents the coding sequence of human B7-L_h2
 XX

SQ Sequence 1168 BP; 299 A; 302 C; 288 G; 279 T; 0 U; 0 Other;

Query Match 78.7%; Score 976; DB 6; Length 1168;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 10%; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 163 AGTCCTGAAGGGCTCCAGGCTCGCTTCAACTGCACCGCTCCAGGGCTGGAAGTCTCAT 222

Db 91 AGTCCTGAAGGGCTCCAGGCTCGCTTCAACTGCACCGCTCCAGGGCTGGAAGTCTCAT 150

QY 223 CATGTGGGCTCTCAGTGACATGTGTGCTTAAGCGTCAGGCCCATGAGGCCCATCATCAC 282

Db 151 CATGTGGGCTCTCAGTGACATGTGTGCTTAAGCGTCAGGCCCATGAGGCCCATCATCAC 210

QY 283 CAATGACCGCTTCACCTCTCAGAGGTACGACAGGGGGGAACTTCACTCGGAGATGAT 342

Db 211 CAATGACCGCTTCACCTCTCAGAGGTACGACAGGGGGGAACTTCACTCGGAGATGAT 270

QY 343 CATCCACAATGTGGAGCCCGAGTGATTCGGGGAAACATCAGATGACGCTCCAGCAACAGTCG 402

Db 271 CATCCACAATGTGGAGCCCGAGTGATTCGGGGAAACATCAGATGACGCTCCAGCAACAGTCG 330

QY 403 CTTGCATGGATCTGCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTTATCCAGTGT 462

Db 331 CTTGCATGGATCTGCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTTATCCAGTGT 390

QY 463 TAATCTTGTAGTCGCTGAGAATGAACCTTGTGAAGTTACTTGTCTACCTCAGCTGGAC 522

Db 391 TAATCTTGTAGTCGCTGAGAATGAACCTTGTGAAGTTACTTGTCTACCTCAGCTGGAC 450

QY 523 CCGGCTCCCGGATATTTCTCTGGAGCTCGGTCTCTCTGGTCAGGCATTCAAGCTATTATTT 582

Db 451 CTGGCTCCCGGATATTTCTCTGGAGCTCGGTCTCTCTGGTCAGGCATTCAAGCTATTATTT 510

QY 583 TGTTCGGAGCCCGAGCGACCTTCAAAGTGCACTGAGCATCTCTGGCTCTGACCCACAGAG 642

Db 511 TGTTCGGAGCCCGAGCGACCTTCAAAGTGCACTGAGCATCTCTGGCTCTGACCCACAGAG 570

QY 643 CAATGGGACTTTGACATTTGGTGGCTACCTCGGAGAGCCCTGAAGCCCGGAAGTCTGCAAC 702

Db 571 CAATGGGACTTTGACATTTGGTGGCTACCTCGGAGAGCCCTGAAGCCCGGAAGTCTGCAAC 630

QY 703 TGTAAATCTCACCTGTGATTCGGTGTCCCAAGACACTGGAGGTGGTATTAAATATTCAGG 762

Db 631 TGTAAATCTCACCTGTGATTCGGTGTCCCAAGACACTGGAGGTGGTATTAAATATTCAGG 690

QY 763 TGTATTATCAAGTTTACCGAGTTTAGTGTTCCTTCACTTGGGCGCAAGTGTGACT 822

Db 691 TGTATTATCAAGTTTACCGAGTTTAGTGTTCCTTCACTTGGGCGCAAGTGTGACT 750

Qy	823	TGACNCTAGCAGGCACCATGCTTCTGACGCCGCGAGTGTTACTCTTACAACTACGCTGCTGCTG	882
Db	751	TGGCACTACGAGGCACCATGCTTCTGACGCCGCGAGTGTTACTCTTACAACTACGCTGCTGCTG	810
Qy	883	CTGCGCCGCTGCTGTGTGTGCGTCSGAACCTGCTGCTGCGGTGTGTTTCTGCTGTGTAGAAG	942
Db	811	CTGCCGGCGTCGTTGTTGTGGCTGCAACTGCTGCTGCGGTGTGTTTCTGCTGTGTAGAAG	870
Qy	943	AAAAAGAGGATTTCGTGATTCAAATTTCAAAAGAAATCTGAAAAGAGAGAAGACAACAAGA	1002
Db	871	AAAAAGAGGATTTCGTGATTCAAATTTCAAAAGAAATCTGAAAAGAGAGAAGACAACAAGA	930
Qy	1003	AACGTGACAGAGAAAGTGGAATAAGAAACTCGGGCTACAAATTCAGATGAAACAAAGACCAC	1062
Db	931	AACCTGACAGAGAAAGTGGAATAAGAAACTCGGGCTACAAATTCAGATGAAACAAAGACCAC	990
Qy	1063	AGAAACCGCTTCTCTCCCTCCCRAATCCTGTGAATCCAGTGATCCTCTGAAACAAAGAACAG	1122
Db	991	AGACCCGCTTCTCTCCCTCCCRAATCCTGTGAATCCAGTGATCCTCTGAAACAAAGAACAG	1050
Qy	1123	TAGCTGTGGCCCTCCTCACGAGCGGGTGATCAACGTCCTCCAGGCCAGGCAAGTCATCC	1182
Db	1051	TAGCTGTGGCCCTCCTCACGAGCGGGTGATCAACGTCCTCCAGGCCAGGCAAGTCATCC	1110
Qy	1183	ACAGGCTTCTTTTAATCTGGCCAGTCTGTGAGAGGTCAGTAATACAACTGTAGTATAG	1240
Db	1111	ACAGGCTTCTTTTAATCTGGCCAGTCTGTGAGAGGTCAGTAATACAACTGTAGTATAG	1168

RESULT 5

RESOLUT 3
ABK13031

ID ABK13031

XX
ID 56K15057 56A10010, 56N11, 56N12

AC ABK13031;

XX	
DT	23-APR-2002 (first entry)
XX	
XX	DNA encoding human B7-like protein, B7-L_h4.
DE	
XX	
XX	Human; B7-Like protein; B7-L; antiinfertility; gynaecological;
KW	antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;
KW	antiinflammatory; dermatological; antipsoriatic; neuroprotective;
KW	antidiabetic; haemostatic; antithyroid; antilulcer; antiallergic;
KW	antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;
KW	reproductive disorder; graft versus host disease; autoimmune disease;
KW	toxic shock syndrome; allergy; nephropathy; skin disorder;
KW	endocrinopathy; lymphoproliferative disorder; gene; ss.

OS Homo sapiens.

CS	nomo sapiens.	Location/Qualifiers
XX		1. .1134
FH	Key	
FT	CDS	

2
2
)

4 E
4 E

4 E
4 E

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

PN WO200200710-A2

XX
FN
W0200Z00710-ARZ.

03-JAN-2002

FD
yy
03-JAN-2002.

XX
DE
28-TIM-3001-3001WA-

PF	28-JUN-2001; 2001WO-US020719.
XX	
XX	
PR	28-JUN-2000; 2000US-0214512P.
PR	28-NOV-2000; 2000US-00729264.
XX	
XX	(AMGE-) AMGEN INC.
PA	
XX	
XX	Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
PI	
XX	
XX	WPI; 2002-130881/17.
DR	
XX	
PT	New B7-like polypeptides, polynucleotides and their modulators, useful
PT	for diagnosing, preventing and treating reproductive, immune and
PT	proliferative disorders, e.g. cancer and arteriosclerosis.
PT	

Claim 1: Fig 4: 13500: English.

The invention relates to an isolated B7-like (B7-L) polypeptide (I). The polypeptide, polynucleotide encoding it and antibody against (I) are useful for treating B7-like polypeptide-related disease, disorders or conditions including reproductive disorders (e.g. infertility), miscarriage, preterm labour and delivery and endometriosis) and proliferative disorders. Antibodies, soluble proteins comprising extracellular domains and other regulators of B7-L polypeptides are useful for enhancing the immune response to tumours. (I) plays a role in growth and maintenance of cancer cells based on the observation of seminal vesicle hyperplasia in transgenic mice overexpressing B7-L polypeptide. Hence modulators of (I) are useful for the treatment of cancer including seminal vesicle cancer, lung, brain, breast, ovarian, testicular cancer and cancers of haematopoietic system. B7-L polypeptide pathway can be manipulated to regulate cytotoxic T-lymphocyte response in allograft transplantation, graft versus host disease, T-cell dependent B cell mediated diseases and autoimmune diseases. B7-L molecules are useful for alleviating the symptoms associated with diseases involving chronic immune cell dysfunction or to treat autoimmune diseases such as systemic lupus erythematosus, rheumatoid arthritis, multiple sclerosis, diabetes, immune thrombocytopenic purpura and psoriasis, chronic inflammatory disease such as inflammatory bowel disease (Crohn's disease and ulcerative colitis), Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They are also useful as immunosuppressive agents for bone marrow and organ transplantation or to prolong graft survival. B7-L molecules are also useful for diagnosis and treatment of diseases involving abnormal cell proliferation, including arteriosclerosis and vascular restenosis. Antagonists of B7-L polypeptides are useful for alleviation of toxic shock syndrome or allo sensitisation due to blood transfusions, and for treatment of allergy, asthma and hypersensitivity reactions, nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and pemphigoid), endocrinopathies (Grave's disease), various pneumopathies (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia, thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and lymphoproliferative disorders such as multiple myeloma. The present sequence represents the coding sequence of human B7-L h4.

Sequence 1139 BP: 290 A: 300 C: 283 G: 266 T: 0 U: 0 Other:

Query Match

Query Match Best Local Similarity 39.6%; SCORE 739; Pred No. 0:

Best Local Similarity 99.9%, Freq: NO. 0,
Matches 789: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

1 63 A C T C C T G C A A C C C C T C C C A C C C T C C C A C C T C A T 222

QY 163 AGTCCGGAAGGGCTCCAGGGCTGGCTTCAACTGCACTGCTCCAGGGCTGGAAAGCTCAT 222

Db 84 AGTCCTGAAGGGCTCCAGGCTCGCTTCAACTGCACCGTCTCCAGGGCTGGAAGCTCAT 143

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
[
\
]
^
_
`
a
b
c
d
e
f
g
h
i
j
k
l
m
n
o
p
q
r
s
t
u
v
w
x
y
z
{
|
}
~
_

D_b 144 CATGTGGGCTCTCAGTGACATGGTGTGCTAAGCGTCAGGCCCATGGAGCCCATCATCAC 203

QY 283 CAATGACCGCTTCACCTCTCAGAGGTACGACCAGGCGGGAACCTTCACCTCGGAGATGAT 342

[illegible]

343 CATCCACAATGTGGAGCCAGTGATTCGGGGAACATCAGATGCAGGCCTCCAGAACAGTCG 402

323
 322
 321
 320
 319
 318
 317
 316
 315
 314
 313
 312
 311
 310
 309
 308
 307
 306
 305
 304
 303
 302
 301
 300
 299
 298
 297
 296
 295
 294
 293
 292
 291
 290
 289
 288
 287
 286
 285
 284
 283
 282
 281
 280
 279
 278
 277
 276
 275
 274
 273
 272
 271
 270
 269
 268
 267
 266
 265
 264
 263
 262
 261
 260
 259
 258
 257
 256
 255
 254
 253
 252
 251
 250
 249
 248
 247
 246
 245
 244
 243
 242
 241
 240
 239
 238
 237
 236
 235
 234
 233
 232
 231
 230
 229
 228
 227
 226
 225
 224
 223
 222
 221
 220
 219
 218
 217
 216
 215
 214
 213
 212
 211
 210
 209
 208
 207
 206
 205
 204
 203
 202
 201
 200
 199
 198
 197
 196
 195
 194
 193
 192
 191
 190
 189
 188
 187
 186
 185
 184
 183
 182
 181
 180
 179
 178
 177
 176
 175
 174
 173
 172
 171
 170
 169
 168
 167
 166
 165
 164
 163
 162
 161
 160
 159
 158
 157
 156
 155
 154
 153
 152
 151
 150
 149
 148
 147
 146
 145
 144
 143
 142
 141
 140
 139
 138
 137
 136
 135
 134
 133
 132
 131
 130
 129
 128
 127
 126
 125
 124
 123
 122
 121
 120
 119
 118
 117
 116
 115
 114
 113
 112
 111
 110
 109
 108
 107
 106
 105
 104
 103
 102
 101
 100
 99
 98
 97
 96
 95
 94
 93
 92
 91
 90
 89
 88
 87
 86
 85
 84
 83
 82
 81
 80
 79
 78
 77
 76
 75
 74
 73
 72
 71
 70
 69
 68
 67
 66
 65
 64
 63
 62
 61
 60
 59
 58
 57
 56
 55
 54
 53
 52
 51
 50
 49
 48
 47
 46
 45
 44
 43
 42
 41
 40
 39
 38
 37
 36
 35
 34
 33
 32
 31
 30
 29
 28
 27
 26
 25
 24
 23
 22
 21
 20
 19
 18
 17
 16
 15
 14
 13
 12
 11
 10
 9
 8
 7
 6
 5
 4
 3
 2
 1
 0

264 CATCCACAATGTGGAGCCAGTGATTCGGGAACATCAGATGCAGCTTCAGAACAGTCC 323

403 CCTGCATGGATCTCTGCTTACCTTACCGTCCAAGTTATGGGAGAGCTGTTCAATCCCAAGTGT 462

[illegible]


```

RESULT 7
ACH16130
ID ACH16130 standard; cDNA; 474 BP.
XX
XX ACH16130;
DT 13-OCT-2003 (first entry)
XX
DE Human adult heart cDNA #444.
XX
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.
XX
XX Homo sapiens.
XX
XX US2003073623-A1.
XX
XX 17-APR-2003.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX (DRMA/) DRMANAC R T.
XX (LABA/) LABAT I.
XX (STAC/) STACHE-CRAIN B.
XX (DICK/) DICKSON M C.
XX (JONE/) JONES L W.
XX
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 3342; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for
XX identifying expressed genes or for physical mapping of the human genome,
XX in forensics, in assessing biodiversity, or in identifying mutations
XX responsible for genetic disorders and other traits. The nucleotide
XX sequences are also useful as hybridisation probes, as oligomers for PCR,
XX for chromosome and gene mapping, in the recombinant production of
XX protein, or in generating antisense DNA or RNA. The purified polypeptide
XX is useful for generating antibodies specific for it. The present sequence
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
XX Sequence 474 BP; 154 A; 128 C; 96 G; 95 T; 0 U; 1 Other;

Query Match 21.6%; Score 268; DB 8; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.4e-124;
Matches 268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 973 GAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGGAAATGAAACTC 1032
DB 63 GAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGGAAATGAAACTC 122
QY 1033 CGGCTCAATTGAGATGAACAAAAGACACAGAAACCGCTTCTCCCTCCCAATCCCTG 1092
DB 123 CGGCTCAATTGAGATGAACAAAAGACACAGAAACCGCTTCTCTCTCCCAATCCCTG 182

RESULT 8
AAI36582
ID AAI36582 standard; DNA; 401 BP.
XX
XX AAI36582;
DT 17-OCT-2001 (first entry)
XX
XX Probe #5268 used to measure gene expression in human placenta sample.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
XX
XX Claim 25; SEQ ID NO 5268; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
XX The present sequence is one such probe. The probes are useful for
XX producing a microarray for predicting, measuring and displaying gene
XX expression in samples derived from human placenta. The probes are useful
XX for antenatal diagnosis of human genetic disorders
XX
XX Sequence 401 BP; 105 A; 98 C; 111 G; 87 T; 0 U; 0 Other;

Query Match 11.6%; Score 144; DB 4; Length 401;
Best Local Similarity 100.0%; Pred. No. 5.7e-62;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 AGTCCTGAAGGGCTCCAGGCTCGCTTCAACTGCACCGTCTCCAGGGCTGGAGCTCAT 222
DB 258 AGTCCTGAAGGGCTCCAGGCTCGCTTCAACTGCACCGTCTCCAGGGCTGGAGCTCAT 317
QY 223 CATGTGGGTCTTCAGTGACATGGTGTGCTTAAGCGTTCAGGCCCATGAGCCCATCATCAC 282
DB 318 CATGTGGGTCTTCAGTGACATGGTGTGCTTAAGCGTTCAGGCCCATGAGCCCATCATCAC 377

```

```
QY 283 CAATGACCGCTTCACTCTCAGAG 306
DB 378 CAATGACCGCTTCACTCTCAGAG 401

RESULT 9
AAI04335
ID AAI04335 standard; DNA; 357 BP.
XX
AC AAI04335;
XX
DT 09-OCT-2001 (first entry)
XX
DE Probe #4326 used to measure gene expression in human breast sample.
XX
KW Probe; human; breast disease; breast cancer; development disorder; ss;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US000661.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression in
PT a human breast.
XX
PS Claim 25; SEQ ID NO 4326; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and non-
CC carcinoma tumours. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pcr_sequences
XX
SQ Sequence 357 BP; 117 A; 68 C; 64 G; 108 T; 0 U; 0 Other;

Query Match 2.0%; Score 25; DB 5; Length 357;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 949 AGGATTTCGATTCAATTCAAAG 973
DB 42 AGGATTTCGATTCAATTCAAAG 66

RESULT 10
ABLI5493/c
ID ABLI5493 standard; cDNA; 6507 BP.
```

```
XX ABL15493;
AC
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 40961.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PR P-PSDB; ABB71390.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 40961; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
CC sequences (ABLI01840-ABLI6175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcr_sequences
XX
SQ Sequence 6507 BP; 1745 A; 1888 C; 1712 G; 1162 T; 0 U; 0 Other;

Query Match 1.9%; Score 24; DB 4; Length 6507;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 872 CGCTGCTGCTGCTGCTGCTGCTGCT 895
DB 2946 CGCTGCTGCTGCTGCTGCTGCTGCT 2923

RESULT 11
ABLI5492/c
ID ABLI5492 standard; cDNA; 59967 BP.
XX
AC ABLI5492;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 40958.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
```

```

PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX P-PSDB; ABB711389.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX
XX Claim 1; SEQ ID NO 40958; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ffp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 59967 BP; 17305 A; 12463 C; 12291 G; 17908 T; 0 U; 0 Other;
XX
XX Query Match 1.9%; Score 24; DB 4; Length 59967;
XX Best Local Similarity 100.0%; Pred. No. 0.25;
XX Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 872 CGTGTGCTGCTGCGCGCGTGT 895
XX
XX 4596 CGCTGCTGCTGCTGCGCGCGTGT 4573
XX
XX
XX
XX RESULT 12
XX ID ABL13032
XX ABK13032 standard; cDNA; 1195 BP.
XX
XX AC ABL13032;
XX
XX 23-APR-2002 (first entry)
XX
XX DNA encoding mouse B7-like protein, B7-L_m1.
XX
XX Mouse; B7-like protein; B7-L; antiinfertility; gynaecological;
XX antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;
XX antiinflammatory; dermatological; antipsoriatic; neuroprotective;
XX antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;
XX antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;
XX reproductive disorder; graft versus host disease; autoimmune disease;
XX toxic shock syndrome; allergy; nephropathy; skin disorder;
XX endocrinopathy; lymphoproliferative disorder; gene; ss.
XX
XX Mus musculus.
XX
XX Key Location/Qualifiers
XX CDS 53..1165
XX /*tag= a
XX /product= "B7-like protein, B7-L_m1"
XX
XX WO200200710-A2.
XX
XX 03-JAN-2002.
XX
XX 28-JUN-2001; 2001WO-US020719.

```

```

XX
XX 28-JUN-2000; 2000US-0214512P.
XX 28-NOV-2000; 2000US-00729264.
XX
XX (AMGE-) AMGEN INC.
XX
XX Welcher AA, Sarmiento UM, Schultz HU, Chute HT;
XX
XX WPI; 2002-130881/17.
XX P-PSDB; AAU75544.
XX
XX New B7-like polypeptides, polynucleotides and their modulators, useful
XX for diagnosing, preventing and treating reproductive, immune and
XX proliferative disorders, e.g. cancer and arteriosclerosis.
XX
XX Claim 1; Fig 5; 135pp; English.
XX
XX The invention relates to an isolated B7-like (B7-L) polypeptide (I). The
XX polypeptide, polynucleotide encoding it and antibody against (I) are
XX useful for treating B7-like polypeptide-related disease, disorders or
XX conditions including reproductive disorders (e.g. infertility,
XX miscarriage, preterm labour and delivery and endometriosis) and
XX proliferative disorders. Antibodies, soluble proteins comprising
XX extracellular domains and other regulators of B7-L polypeptides are
XX useful for enhancing the immune response to tumours. (I) plays a role in
XX growth and maintenance of cancer cells based on the observation of
XX seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
XX polypeptide. Hence modulators of (I) are useful for the treatment of
XX cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
XX testicular cancer and cancers of haematopoietic system. B7-L polypeptide
XX pathway can be manipulated to regulate cytotoxic T-lymphocyte response in
XX allograft transplantation, graft versus host disease, T-cell dependent B-
XX cell mediated diseases and autoimmune diseases. B7-L molecules are useful
XX for alleviating the symptoms associated with diseases involving chronic
XX immune cell dysfunction or to treat autoimmune diseases such as systemic
XX lupus erythematosus, rheumatoid arthritis, multiple sclerosis, diabetes,
XX immune thrombocytopenic purpura and psoriasis, chronic inflammatory
XX disease such as inflammatory bowel disease (Crohn's disease and
XX ulcerative colitis), Grave's disease, Hashimoto's thyroiditis and
XX diabetes mellitus. They are also useful as immunosuppressive agents for
XX bone marrow and organ transplantation or to prolong graft survival. B7-L
XX molecules are also useful for diagnosis and treatment of diseases
XX involving abnormal cell proliferation, including arteriosclerosis and
XX vascular restenosis. Antagonists of B7-L polypeptides are useful for
XX alleviation of toxic shock syndrome or allosensitisation due to blood
XX transfusions, and for treatment of allergy, asthma and hypersensitivity
XX reactions, nephropathies (e.g. glomerulonephritis), skin disorders
XX (pemphigus and pemphigoid), endocrinopathies (Grave's disease), various
XX pneumopathies (extrinsic alveolitis), vasculopathies, coeliac disease,
XX anaemia, thrombocytopenias, Guillain-Barre syndrome and myasthenia
XX gravis, and lymphoproliferative disorders such as multiple myeloma. The
XX present sequence represents the coding sequence of mouse B7-L_m1
XX
XX Sequence 1195 BP; 313 A; 304 C; 291 G; 287 T; 0 U; 0 Other;
XX
XX Query Match 1.9%; Score 23; DB 6; Length 1195;
XX Best Local Similarity 100.0%; Pred. No. 0.73;
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 527 CTCCTCGGATATTCTCTGGAGCT 549
XX
XX 521 CTCCTCGGATATTCTCTGGAGCT 543
XX
XX
XX RESULT 13
XX AAV99359
XX ID AAV99359 standard; DNA; 141 BP.
XX
XX AC AAV99359;
XX
XX 25-MAR-1999 (first entry)
XX
XX 5' PCR primer for GDNFR-gammal cDNA encoding the full length protein.

```

XX Glial cell line-derived neurotrophic factor receptor gamma 1;
 KW GDNFR-alpha; glial cell line-derived neurotrophic factor; GDNF;
 KW neurodegenerative disease; anyotrophic lateral sclerosis; GDNFR-gamma1;
 KW Parkinson's disease; schizophrenia; insomnia; tardive dyskinesia;
 KW hypertension; pituitary adenomas; hyperprolactinemia; thyroid tumour;
 KW renal disorder; kidney failure; gut dysfunction; regeneration;
 KW cardiomyocyte; epithelium; hepatocyte; PCR primer; ss.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX WO9853069-A2.
 XX
 PD 26-NOV-1998.
 XX
 XX 20-MAY-1998; 98WO-US010328.
 PF
 XX 20-MAY-1997; 97US-0047092P.
 PR
 XX 27-JUN-1997; 97US-00884638.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 XX Ni J, Hsu T, Young P, Gentz RL, Ruben SM;
 PI WPI; 1999-070150/06.
 XX
 DR New isolated glial cell derived neurotrophic factor receptors - used to
 PT develop products for treating e.g. neurodegenerative disorders,
 PT schizophrenia, hypertension, tumours, renal disorders, kidney failure or
 PT gut dysfunction.
 XX
 XX Example 6; Page 68; 156pp; English.
 PS
 XX PCR primers AAV99359-60 were used to amplify cDNA encoding the full
 CC length protein of a glial cell line-derived neurotrophic factor receptor
 CC gamma 1 (GDNFR-gamma1). The amplified product was subsequently cloned and
 CC expressed in Baculovirus. GDNFR-beta shares high homology with GDNFR-
 CC alpha, which is capable of complexing with glial cell line-derived
 CC neurotrophic factor (GDNF) and mediating cell response to GDNF. The GDNFR
 CC polypeptides and agonists can be used for treating disorders associated
 CC with decreased activity of the respective polypeptides. They can be used
 CC for treating neurodegenerative diseases such as amyotrophic lateral
 CC sclerosis, Parkinson's disease, schizophrenia, insomnia, tardive
 CC dyskinesia, hypertension, pituitary adenomas, hyperprolactinemia, thyroid
 CC tumour, renal disorders, kidney failure, gut dysfunction, or for
 CC regeneration of cardiomyocytes, epithelium or hepatocytes. Antagonists of
 CC the polypeptides can be used for treating disorders associated with
 CC increased activity of the respective polypeptides. The products can also
 CC be used for detection, diagnosis and drug screening
 XX
 SQ Sequence 141 BP; 19 A; 60 C; 38 G; 24 T; 0 U; 0 Other;
 Query Match 1.8%; Score 22; DB 2; Length 141;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 873 GCTGCTGCTGCTCCGCCGCG 894
 Db 69 GCTGCTGCTGCTCCGCCGCG 90
 RESULT 14
 ABN99162
 ID ABN99162 standard; DNA; 767 BP.
 XX
 AC ABN99162;
 XX
 DT 01-AUG-2002 (first entry)
 XX
 DE Arabidopsis thaliana expressed polynucleotide SEQ ID NO 930.
 XX
 XX Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;
 KW

KW disease; crop; thale cress; tolerance factor; insect; pathogen;
 KW nutrition; ds.
 XX
 OS Arabidopsis thaliana.
 PN US2002023281-A1.
 XX
 PD 21-FEB-2002.
 XX
 PF 26-JAN-2001; 2001US-00770445.
 XX
 PR 27-JAN-2000; 2000US-0178472P.
 XX
 XX (GORL/) GORLACH J.
 PA (ANYV/) AN Y.
 PA (HAML/) HAMILTON C M.
 PA (PRIC/) PRICE J L.
 PA (RAIN/) RAINES T M.
 PA (YUYV/) YU Y.
 PA (RAME/) RAMEAKA J G.
 PA (PAGE/) PAGE A.
 PA (MATH/) MATHW A V.
 PA (LEDF/) LEDFORD B L.
 PA (WOES/) WOESSNER J P.
 PA (HAAS/) HAAS W D.
 PA (GARC/) GARCIA C A.
 PA (KRICK/) KRICKER M.
 PA (SLAT/) SLATER T.
 PA (DAVI/) DAVIS K R.
 PA (ALLE/) ALLEN K.
 PA (HOFF/) HOFFMAN N.
 PA (HURB/) HURBAN P.
 XX
 XX Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
 PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
 PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
 XX Hurban P;
 XX WPI; 2002-403163/43.
 XX New Arabidopsis thaliana nucleic acid for identifying homologous genes,
 PT producing compositions that modulate the expression or function of its
 PT encoded protein, and mapping functional regions of protein.
 XX
 PS Claim 1; SEQ ID NO 930; 49pp + Sequence Listing; English.
 XX The invention relates to an Arabidopsis thaliana nucleic acid (I)
 CC comprising a sequence capable of hybridising under stringent conditions
 CC to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),
 CC given in the specification or its fragment. A polypeptide (II) encoded by
 CC (I), a transgenic plant (III) comprising an exogenous nucleic acid or a
 CC genetically modified cell (IV) comprising an exogenous nucleic acid, is
 CC useful for screening a candidate agent for its biological effect. (I) is
 CC useful in identifying homologous or related genes, in producing
 CC compositions that modulate the expression or function of its encoded
 CC protein, mapping functional regions of the protein and in studying
 CC associated physiological pathways. (I) is also useful for the genetic
 CC manipulation of cells, particularly plant cells. (I) is also useful in
 CC screening assays of various plant strains to determine the strains that
 CC are best capable of withstanding a particular disease or environmental
 CC stress. (II) and (III) are useful for screening of biologically active
 CC agents, e.g. fungicides, insecticides, etc., for elucidating biochemical
 CC pathways. The screened agents are useful in improved methods of treating
 CC crops to prevent or treat disease. (II) are also useful in screening
 CC programs to identify agents that mimic or enhance the action of tolerance
 CC factors. Such agents are useful in improved methods of treating crops to
 CC enhance their tolerance to environmental stress. (I) is also useful for
 CC enhancing or inhibiting production of a biosynthetic product in a plant.
 CC (III) is useful for identifying other mediators that may induce
 CC expression of proteins of interest, for establishing the extent to which
 CC any specific insect and/or pathogen is responsible for damage to a
 CC particular plant, for identifying other mediators that enhance or induce
 CC tolerance to environmental stress, for identifying factors involved in

CC biosynthetic pathways of nutritional, commercial, or medicinal value and
 CC for identifying productions of nutritional, commercial or medicinal
 CC value. (IV) is useful in the study of genetic function and regulation,
 CC for alteration of the cellular metabolism and for screening compounds
 CC that may affect the biological function of the gene or gene products.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html?DocID=999909770445
 XX
 SQ Sequence 767 BP; 169 A; 199 C; 147 G; 252 T; 0 U; 0 Other;
 Query Match 1.8%; Score 22; DB 6; Length 767;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 886 CCGCGCTCGTTGTTGGCTGC 907
 DB 672 CCGCGCTCGTTGTTGGCTGC 693

RESULT 15
 AAV35364
 ID AAV35364 standard; cDNA; 1200 BP.
 XX
 AC AAV35364;
 XX
 DT 28-SEP-1998 (first entry)
 XX
 DE Human GDNF alpha-3 receptor cDNA #1.
 XX
 KW Glial cell line-derived neurotrophic factor alpha-3 receptor; GDNF;
 KW treatment; neurodegenerative disease; Parkinson's Disease; ALS; SMA;
 KW amyotrophic lateral sclerosis; spinal muscular atrophy; nerve; trauma;
 KW Huntington's Disease; Alzheimer's Disease; diabetic neuropathy; muscle;
 KW muscular dystrophy; diagnostic; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT CDS
 FT 1..1200
 FT /*tag= a
 FT /product= "GDNF alpha-3"
 FT /notes= "partial sequence of glial cell-derived
 FT neurotrophic factor alpha-3 receptor"
 XX
 PN EP846764-A2.
 XX
 PD 10-JUN-1998.
 XX
 PF 20-NOV-1997; 97EP-00309375.
 XX
 PR 27-NOV-1996; 96GB-00024677.
 PR 09-MAY-1997; 97GB-00009463.
 XX
 PA (SMIX) SMITHKLINE BEECHAM PLC.
 XX
 PI Lawrence GMP;
 XX
 DR WPI; 1998-299980/27.
 DR F-PSDB; AAW65116.
 XX
 PT New factor alpha 3 receptor polypeptide and e.g. DNA and agonists - used
 PT to treat neuro degenerative diseases, muscular diseases and nerve and
 PT muscle trauma and in diagnostic assays.
 XX
 PS Claim 9; Fig 1; 22pp; English.
 XX
 CC This sequence encodes a novel glial cell line-derived neurotrophic factor
 CC alpha-3 receptor (GDNF alpha-3). This protein can be used to treat e.g.
 CC neurodegenerative diseases (such as Parkinson's Disease, amyotrophic
 CC lateral sclerosis (ALS), spinal muscular atrophy (SMA), Huntington's
 CC Disease, Alzheimer's Disease, diabetic neuropathy), muscular diseases
 CC (including the muscular dystrophies) and nerve and muscle trauma and in

CC diagnostic assays for such conditions
 XX
 SQ Sequence 1200 BP; 236 A; 419 C; 307 G; 238 T; 0 U; 0 Other;
 Query Match 1.8%; Score 22; DB 2; Length 1200;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 873 GCTGCTGCTGTCGCCGCTGC 894
 DB 51 GCTGCTGCTGTCGCCGCTGC 72

Search completed: September 14, 2004, 23:41:50
 Job time : 533.923 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 23:14:24 ; Search time 96.902 Seconds
(without alignments)
7101.389 Million cell updates/sec

Title: US-09-729-264-5

Perfect score: 1240

Sequence: 1 agdgtgagtcagcccaaca.....gtaatacaactgtagtatag 1240

Scoring table: OLIGO NUC

Gapop_60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1281558

Minimum DB seq length: 16

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
- 2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
- 3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
- 4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
- 5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	1.8	1203	US-09-220-528-65	Sequence 65, Appl
2	22	1.8	1699	US-09-187-906-20	Sequence 20, Appl
3	22	1.8	3942	US-09-162-484-19	Sequence 19, Appl
C 4	21	1.7	550	US-09-669-751-128	Sequence 128, App
C 5	21	1.7	1428	US-09-489-039A-3243	Sequence 3243, App
C 6	20	1.6	394	US-09-621-976-16752	Sequence 16752, A
7	20	1.6	794	US-09-621-976-16751	Sequence 16751, A
8	20	1.6	984	US-09-252-991A-12897	Sequence 12897, A
9	20	1.6	1857	US-09-252-991A-12735	Sequence 12735, A
C 10	20	1.6	1910	US-09-593-711A-3	Sequence 3, Appli
C 11	20	1.6	1914	US-07-601-094-1	Sequence 1, Appli
C 12	20	1.6	1914	US-08-012-735-1	Sequence 1, Appli
13	20	1.6	2214	US-08-864-038A-1	Sequence 1, Appli
14	20	1.6	3331	US-08-864-038A-2	Sequence 2, Appli
15	20	1.6	3331	US-08-864-038A-4	Sequence 4, Appli
16	20	1.6	6407	US-08-616-844-7	Sequence 7, Appli
17	20	1.6	6407	US-08-599-654-7	Sequence 7, Appli
18	20	1.6	6407	US-08-944-868A-7	Sequence 7, Appli
19	20	1.6	6407	US-08-944-423A-7	Sequence 7, Appli
20	20	1.6	6407	US-08-944-496-7	Sequence 7, Appli
C 21	19	1.5	402	US-09-621-976-88	Sequence 88, Appl
C 22	19	1.5	1425	US-08-464-148-1	Sequence 1, Appli
C 23	19	1.5	1425	US-08-385-500-1	Sequence 1, Appli
C 24	19	1.5	1425	US-08-846-784-1	Sequence 1, Appli
C 25	19	1.5	1477	US-09-620-312B-1019	Sequence 1019, App
C 26	19	1.5	1768	US-09-833-381-523	Sequence 523, App
27	19	1.5	1899	US-09-919-060-15	Sequence 15, Appl

C 28	19	1.5	1899	4	US-09-919-060-16	Sequence 16, Appl
29	19	1.5	1989	1	US-08-447-500-5	Sequence 5, Appli
30	19	1.5	1989	1	US-08-454-097-5	Sequence 5, Appli
31	19	1.5	1989	1	US-08-453-866-5	Sequence 5, Appli
32	19	1.5	1989	3	US-08-185-359-5	Sequence 5, Appli
33	19	1.5	1995	4	US-09-620-312D-908	Sequence 908, App
34	19	1.5	2023	4	US-09-491-522-6	Sequence 6, Appli
C 35	19	1.5	2115	2	US-08-474-379C-60	Sequence 60, Appl
C 36	19	1.5	2115	3	US-09-146-249A-60	Sequence 60, Appl
C 37	19	1.5	2115	3	US-08-206-188B-60	Sequence 60, Appl
38	19	1.5	2439	4	US-09-489-039A-4904	Sequence 4904, Ap
39	19	1.5	2450	4	US-09-491-522-2	Sequence 2, Appli
40	19	1.5	2617	4	US-09-786-240-21	Sequence 21, Appl
41	19	1.5	2693	4	US-09-919-060-12	Sequence 12, Appl
C 42	19	1.5	2693	4	US-09-919-060-14	Sequence 14, Appl
C 43	19	1.5	3311	4	US-09-367-891A-5	Sequence 5, Appli
C 44	19	1.5	5092	3	US-09-412-545-1	Sequence 1, Appli
45	19	1.5	5621	4	US-09-566-921-106	Sequence 106, App

ALIGNMENTS

RESULT 1
US-09-220-528-65
; Sequence 65, Application US/09220528A
; Patent No. 6284540
; GENERAL INFORMATION:
; APPLICANT: Milbrandt, Jeffrey D.
; APPLICANT: Baloch, Robert H.
; TITLE OF INVENTION: Artemin, A No. 6284540el Neurotrophic Factor
; FILE REFERENCE: 6029-7998
; CURRENT APPLICATION NUMBER: US/09/220,528A
; CURRENT FILING DATE: 1998-12-24
; EARLIER APPLICATION NUMBER: 09/218,698
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 60/108,148
; EARLIER FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/163,283
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 65
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-528-65

Query Match 1.8%; Score 22; DB 3; Length 1203;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 873 GCTGCTGCTGCTCGCCGCGTCG 894
|||||
Db 51 GCTGCTGCTGCTCGCCGCGTCG 72

RESULT 2
US-09-187-906-20
; Sequence 20, Application US/09187906
; Patent No. 6677135
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: Ret Ligand (RetL) for Stimulating Neural
; TITLE OF INVENTION: and Renal Growth
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02142

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/187,906
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/07726
FILING DATE: 07-MAY-97
APPLICATION NUMBER: US 60/017,427
FILING DATE: 08-MAY-96
APPLICATION NUMBER: US 60/019,300
FILING DATE: 07-JUN-96
APPLICATION NUMBER: US 60/021,859
FILING DATE: 16-JUL-96
APPLICATION NUMBER: US 60/043,533
FILING DATE: 10-APR-97
ATTORNEY/AGENT INFORMATION:
NAME: Kaplan, Warren A.
REGISTRATION NUMBER: 34,199
REFERENCE/DOCKET NUMBER: A008 PCT CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-679-2400
TELEFAX: 617-679-2838
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 1699 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 175..1374
US-09-187-906-20

Query Match 1.8%; Score 22; DB 4; Length 1699;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 873 GCTGCTGCTGCTGCCCGCTCG 894
Db 225 GCTGCTGCTGCTGCCCGCTCG 246

RESULT 3
US-09-162-484-19
; Sequence 19, Application US/09162484
; Patent No. 6248724
; GENERAL INFORMATION:
; APPLICANT: Mohuczy, Dagmara
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS TARGETED TO
; FILE REFERENCE: UFLA:087/UFLA087P
; CURRENT APPLICATION NUMBER: US/09/162,484
; CURRENT FILING DATE: 1998-09-25
; EARLIER APPLICATION NUMBER: 60/059,661
; EARLIER FILING DATE: 1997-09-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 3942
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-162-484-19

Query Match 1.8%; Score 22; DB 3; Length 3942;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 873 GCTGCTGCTGCTGCCCGCTCG 894
Db 72 GCTGCTGCTGCTGCCCGCTCG 93

RESULT 4
US-09-669-751-128/c
; Sequence 128, Application US/09669751
; Patent No. 6551575
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Ralph J.
; TITLE OF INVENTION: Methods for Identifying Compounds for
; TITLE OF INVENTION: Motion Sickness, Vertigo and Other Disorders Related to
; TITLE OF INVENTION: Balance and the Perception of Gravity
; FILE REFERENCE: P-NI 3864
; CURRENT APPLICATION NUMBER: US/09/669,751
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 60/168,579
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Drosophila
US-09-669-751-128

Query Match 1.7%; Score 21; DB 4; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 909 ACTGCTGCTGCCGTGTTGTT 929
Db 317 ACTGCTGCTGCCGTGTTGTT 297

RESULT 5
US-09-489-039A-3243/c
; Sequence 3243, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 3243
; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3243

Query Match 1.7%; Score 21; DB 4; Length 1428;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 186 GCTTCAACTGCACCGTCTCCC 206
Db 1147 GCTTCAACTGCACCGTCTCCC 1127

RESULT 6
US-09-621-976-16752
; Sequence 16752, Application US/09621976
; Patent No. 6639063

```
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16752
; LENGTH: 394
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16752

Query Match      1.6%; Score 20; DB 4; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      882 GCTGCCGCCGCTGCTGTGT 901
        |||||||
Db      34 GCTGCCGCCGCTGCTGTGT 53

RESULT 7
US-09-621-976-16751
; Sequence 16751, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16751
; LENGTH: 794
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16751

Query Match      1.6%; Score 20; DB 4; Length 794;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      882 GCTGCCGCCGCTGCTGTGT 901
        |||||||
Db      34 GCTGCCGCCGCTGCTGTGT 53

US-09-593-711A-3/c
; Sequence 3, Application US/09593711A
; Patent No. 6271030
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Madeline M. Butler
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF C/EBP BETA EXPRESSION
; FILE REFERENCE: RTS-0118
; CURRENT APPLICATION NUMBER: US/09/593,711A
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 244
; SEQ ID NO 3
; LENGTH: 1910
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
; NAME/KEY: unsure
; LOCATION: 1415
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1421
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1422
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1423

US-09-252-991A-12897
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12897

Query Match      1.6%; Score 20; DB 4; Length 984;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      899 TGTGGCTGCAACTGCTGCTG 918
        |||||||
Db      395 TGTGGCTGCAACTGCTGCTG 414

RESULT 9
US-09-252-991A-12735
; Sequence 12735, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12735
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12735

Query Match      1.6%; Score 20; DB 4; Length 1857;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      899 TGTGGCTGCAACTGCTGCTG 918
        |||||||
Db      1746 TGTGGCTGCAACTGCTGCTG 1765

RESULT 10
US-09-593-711A-3/c
; Sequence 3, Application US/09593711A
; Patent No. 6271030
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Madeline M. Butler
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF C/EBP BETA EXPRESSION
; FILE REFERENCE: RTS-0118
; CURRENT APPLICATION NUMBER: US/09/593,711A
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 244
; SEQ ID NO 3
; LENGTH: 1910
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
; NAME/KEY: unsure
; LOCATION: 1415
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1421
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1422
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1423
```

OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 1424
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 1458
OTHER INFORMATION: unknown
NAME/KEY: CDS
LOCATION: (299)...(1336)
US-09-593-711A-3

Query Match 1.6%; Score 20; DB 3; Length 1910;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 872 CGCTGCTGCTGCTGCCGCG 891
Db 205 CGCTGCTGCTGCTGCCGCG 186

RESULT 11
US-07-601-094-1/c
Sequence 1, Application US/07601094
Patent No. 5215892
GENERAL INFORMATION:
APPLICANT: Kishimoto, Tadamitsu
APPLICANT: Hirano, Toshio
APPLICANT: Akira, Shizuo
APPLICANT: Ishiki, Hiroshi
APPLICANT: Tanabe, Osamu
APPLICANT: Kinoshita, Shigemi
APPLICANT: Shimamoto, Takuya
TITLE OF INVENTION: C/BEP2 Gene and Recombinant
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak &
ADDRESS: Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/601,094
FILING DATE: 19901022
CLASSIFICATION: 435
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1914 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 281..1316
OTHER INFORMATION:
US-07-601-094-1

Query Match 1.6%; Score 20; DB 1; Length 1914;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 872 CGCTGCTGCTGCTGCCGCG 891
Db 204 CGCTGCTGCTGCTGCCGCG 185

RESULT 12
US-08-012-735-1/c
Sequence 1, Application US/08012735
Patent No. 5360894
GENERAL INFORMATION:
APPLICANT: Kishimoto, Tadamitsu
APPLICANT: Hirano, Toshio
APPLICANT: Akira, Shizuo
APPLICANT: Ishiki, Hiroshi
APPLICANT: Tanabe, Osamu
APPLICANT: Kinoshita, Shigemi
APPLICANT: Shimamoto, Takuya
TITLE OF INVENTION: C/BEP2 Gene and Recombinant
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak &
ADDRESS: Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/012,735
FILING DATE: 19930203
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/601,094
FILING DATE: 22 OCT 1990
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1914 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 281..1316
OTHER INFORMATION:
US-08-012-735-1

Query Match 1.6%; Score 20; DB 1; Length 1914;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 872 CGCTGCTGCTGCTGCCGCG 891
Db 204 CGCTGCTGCTGCTGCCGCG 185

RESULT 13
US-08-864-038A-1
Sequence 1, Application US/08864038A
Patent No. 6001592
GENERAL INFORMATION:
APPLICANT: Kunio NAKASHIMA et al.
TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE cDNA, VECTOR

```
; TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
; TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
; TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
; TITLE OF INVENTION: TO SAID POLYPEPTIDE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 812-5 Hirano
; STREET: Isshinden
; CITY: Tsu-city
; STATE: Mie-prefecture
; COUNTRY: JAPAN
; ZIP: 514-01
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Microsoft Windows 95
; SOFTWARE: Word Perfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,038A
; FILING DATE: May 28, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-184459
; FILING DATE: 15-July-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-5610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)986-2340
; TELEFAX: (212)953-7733
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2214
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: Pinctada fucata
; CELL TYPE: mantle epithelial cell
; US-08-864-038A-1
;
Query Match 1.6%; Score 20; DB 3; Length 2214;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 872 CGCTGCTGCTGCTGCCGCCG 891
Db 855 CGCTGCTGCTGCTGCCGCCG 874

RESULT 14
US-08-864-038A-2
; Sequence 2, Application US/08864038A
; Patent No. 6001592
; GENERAL INFORMATION:
; APPLICANT: Kunio NAKASHIMA et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
; TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
; TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
; TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 812-5 Hirano
; STREET: Isshinden
; CITY: Tsu-city
; STATE: Mie-prefecture
; COUNTRY: JAPAN
; ZIP: 514-01
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Microsoft Windows 95
; SOFTWARE: Word Perfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,038A
; FILING DATE: May 28, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-184459
; FILING DATE: 15-July-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-5610
; TELECOMMUNICATION INFORMATION:
```

```
; SOFTWARE: Word Perfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,038A
; FILING DATE: May 28, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-184459
; FILING DATE: 15-July-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-5610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)986-2340
; TELEFAX: (212)953-7733
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3331
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Pinctada fucata
; CELL TYPE: mantle epithelial cell
; FEATURE: mRNA
; LOCATION: from 1 to 3331
; IDENTIFICATION METHOD: E (by experiment)
; US-08-864-038A-2
;
Query Match 1.6%; Score 20; DB 3; Length 3331;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 872 CGCTGCTGCTGCTGCCGCCG 891
Db 904 CGCTGCTGCTGCTGCCGCCG 923

RESULT 15
US-08-864-038A-4
; Sequence 4, Application US/08864038A
; Patent No. 6001592
; GENERAL INFORMATION:
; APPLICANT: Kunio NAKASHIMA et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
; TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
; TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
; TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 812-5 Hirano
; STREET: Isshinden
; CITY: Tsu-city
; STATE: Mie-prefecture
; COUNTRY: JAPAN
; ZIP: 514-01
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Microsoft Windows 95
; SOFTWARE: Word Perfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,038A
; FILING DATE: May 28, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-184459
; FILING DATE: 15-July-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-5610
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (212)986-2340
; TELEFAX: (212)953-7733
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3331
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE: Pinctada fucata
; ORGANISM: Pinctada fucata
; CELL TYPE: mantle epithelial cell
; FEATURE:
; NAME/KEY: CDS
; LOCATION: from 50 to 2263
; IDENTIFICATION METHOD: P (by similarity to some other pattern)
US-08-864-038A-4
```

```
Query Match      1.6%; Score 20; DB 3; Length 3331;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      872 CGCTGCTGCTGCTGCCGCCG 891
      |||||
Db      904 CGCTGCTGCTGCTGCCGCCG 923
```

```
Search completed: September 15, 2004, 06:48:27
Job time : 96.902 secs
```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 23:16:34 ; Search time 658.934 Seconds
(without alignments)
9469.459 Million cell updates/sec

Title: US-09-729-264-5
Perfect score: 1240
Sequence: 1 aggtgtgagtcagcgaaca.....gtaatacaactgtatag 1240

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 3304383 seqs, 2515761380 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6509420

Minimum DB seq length: 16

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1125	90.7	2051	16	US-10-104-047-1104
2	268	21.6	474	10	US-09-918-995-3342
3	144	11.6	401	9	US-09-864-761-16305
4	25	2.0	357	9	US-09-864-761-16653
5	24	1.9	1078	13	US-10-027-632-255100
6	24	1.9	1078	13	US-10-027-632-255101
7	24	1.9	1078	13	US-10-027-632-255102
8	24	1.9	1078	16	US-10-027-632-255100
9	24	1.9	1078	16	US-10-027-632-255101
10	24	1.9	1078	16	US-10-027-632-255102
11	24	1.9	1121	13	US-10-027-632-257899
12	24	1.9	1121	16	US-10-027-632-257899
13	23	1.9	1284	17	US-10-437-963-81852
14	22	1.8	767	9	US-09-770-445-930

15	1.8	1203	9	US-09-220-920-65	Sequence 65, Appl
16	1.8	1829	9	US-09-828-366-15	Sequence 15, Appl
17	1.8	4142	12	US-10-152-319A-1498	Sequence 1498, Ap
18	1.8	4571	9	US-09-764-847-1043	Sequence 1043, Ap
19	1.8	4571	15	US-10-092-154-1043	Sequence 1043, Ap
20	1.8	10462	9	US-09-764-847-1044	Sequence 1044, Ap
21	1.8	10462	10	US-09-764-891-6203	Sequence 6203, Ap
22	1.8	10462	15	US-10-092-154-1044	Sequence 1044, Ap
23	1.7	462	17	US-10-437-963-41506	Sequence 41506, A
24	1.7	493	10	US-09-918-995-11051	Sequence 11051, A
25	1.7	550	15	US-10-255-536-128	Sequence 128, App
26	1.7	602	17	US-10-767-701-28842	Sequence 28842, A
27	1.7	606	13	US-10-027-632-134169	Sequence 134169, A
28	1.7	606	16	US-10-027-632-134169	Sequence 134169, A
29	1.7	645	13	US-10-027-632-285369	Sequence 285369, A
30	1.7	645	13	US-10-027-632-285370	Sequence 285370, A
31	1.7	645	16	US-10-027-632-285369	Sequence 285369, A
32	1.7	645	16	US-10-027-632-285370	Sequence 285370, A
33	1.7	1558	10	US-09-930-213-20	Sequence 20, Appl
34	1.7	3362	13	US-10-087-192-185	Sequence 185, App
35	1.7	94917	13	US-10-087-192-184	Sequence 184, App
36	1.7	357652	17	US-10-322-696-34	Sequence 34, Appl
37	1.6	330	17	US-10-437-963-55684	Sequence 55684, A
38	1.6	358	13	US-09-823-245A-58	Sequence 58, Appl
39	1.6	398	9	US-09-963-965-4945	Sequence 4945, Ap
40	1.6	412	10	US-09-918-995-5855	Sequence 5855, Ap
41	1.6	442	13	US-10-027-632-320734	Sequence 320734, A
42	1.6	442	16	US-10-027-632-320734	Sequence 320734, A
43	1.6	507	17	US-10-430-201-970	Sequence 970, App
44	1.6	507	17	US-10-430-201-971	Sequence 971, App
45	1.6	596	17	US-10-767-701-5378	Sequence 5378, Ap

ALIGNMENTS

RESULT 1

US-10-104-047-1104
; Sequence 1104, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392a1el full length cdna
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1104
; LENGTH: 2051
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1104

Query Match	90.7%	Score 1125;	DB 16;	Length 2051;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1125;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	116	GGTTCGGGTCTCGTAAATGAAGTCATAGAGGCCCCAGAGTCAACAGTCTCTGAAGGC	175	
Db	203	GGTTCGGGTCTCGTAAATGAAGTCATAGAGGCCCCAGAGTCAACAGTCTCTGAAGGC	262	
QY	176	TCCAGGCTCGTTCACCTGACCGTCTCCAGGGCTGGAAGCTCATATGTGGGTCTC	235	
Db	263	TCCAGGCTCGTTCACCTGACCGTCTCCAGGGCTGGAAGCTCATATGTGGGTCTC	322	
QY	236	AGTGACATGGTGGTGTAGCGTCAGGCCCATGAGCCCATCATCCATGACCGTTC	295	
Db	323	AGTGACATGGTGGTGTAGCGTCAGGCCCATGAGCCCATCATCCATGACCGTTC	382	
QY	296	ACCTCTCAGAGGTACGACCGAGGGCGGAACTTCACCTCGGAGATGATCATCCCAATGTG	355	

Db 383 ACCTCTCAGAGGTACGACCGGCGGAACTTCCACCTCGAGATGATCATCCACAATGTG 442
QY 356 GAGCCAGGATTCGGGGAACTACAGATCAGCTCCAGACAGTCCGCTGATGATCT 415
Db 443 GAGCCAGGATTCGGGGAACTACAGATCAGCTCCAGACAGTCCGCTGATGATCT 502
QY 416 GCTTACCTTACCGTCAAGTTATGGGAGAGCTGTTTCATTCACAGTGTAAATCTGTAGTC 475
Db 503 GCTTACCTTACCGTCAAGTTATGGGAGAGCTGTTTCATTCACAGTGTAAATCTGTAGTC 562
QY 476 GCTGAGAAATGAACTTGTGAAGTTACTTGTCTACCTCAGCTGAGACCGGCTCCCGAT 535
Db 563 GCTGAGAAATGAACTTGTGAAGTTACTTGTCTACCTCAGCTGAGACCGGCTCCCGAT 622
QY 536 ATTTCCTGGAGCTCGGTCTCCTGCTCAGCCATTCAGCTTATTTTCTCCGAGCCC 595
Db 623 ATTTCCTGGAGCTCGGTCTCCTGCTCAGCCATTCAGCTTATTTTCTCCGAGCCC 682
QY 596 AGCGACCTTCAAAGTSCAGTGAGCATCTGGCTCTGACCCACAGACCAATGGGACTTTG 655
Db 683 AGCGACCTTCAAAGTSCAGTGAGCATCTGGCTCTGACCCACAGACCAATGGGACTTTG 742
QY 656 ACTTGGTGGCTACCTGGAAGAGCTGAAGCCCGCAAGTCTGCAACTGTAATCTCACT 715
Db 743 ACTTGGTGGCTACCTGGAAGAGCTGAAGCCCGCAAGTCTGCAACTGTAATCTCACT 802
QY 716 GTGATTCGGTGTCCCCAAGACACTGGAGGTGGTATTATTAATCCAGGTGATTATCAAGT 775
Db 803 GTGATTCGGTGTCCCCAAGACACTGGAGGTGGTATTATTAATCCAGGTGATTATCAAGT 862
QY 776 TTACCGAGTTAGTTTTTCAITGGCTACTTGGGCAAGTTGGACTTGGACTAGCAGC 835
Db 863 TTACCGAGTTAGTTTTTCAITGGCTACTTGGGCAAGTTGGACTTGGACTAGCAGC 922
QY 836 ACATGCTTCTGACGCGAGCTGTACTCTTACAATAGCTGTCTGTGCGCGCGTCT 895
Db 923 ACCATGCTTCTGACGCGAGCTGTACTCTTACAATAGCTGTCTGTGCGCGCGTCT 982
QY 896 TGTGTGGCTGCAACTGCTGCTGCGCTGTGTTTCTGTGTAGAGAAAGAGGATTT 955
Db 983 TGTGTGGCTGCAACTGCTGCTGCGCTGTGTTTCTGTGTAGAGAAAGAGGATTT 1042
QY 956 CATTATCAATTTCAAAGAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAGAA 1015
Db 1043 CGTATTCAATTTCAAAGAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAGAA 1102
QY 1016 AGTGAATGAAATCTCCGGCTACAAATTCAGATGAACAAAGACCAACAAACCGCTTCT 1075
Db 1103 AGTGAATGAAATCTCCGGCTACAAATTCAGATGAACAAAGACCAACAAACCGCTTCT 1162
QY 1076 CTCCCTCCCAATCTGTGAATCCAGTGATCTGAAACAAAGAAACAGTAGCTGTGCGCT 1135
Db 1163 CTCCCTCCCAATCTGTGAATCCAGTGATCTGAAACAAAGAAACAGTAGCTGTGCGCT 1222
QY 1136 CTTCCACGCGGCTGATCAACGTCCACCCGACGCAAGTCAATCCACAGGCTTCTTTT 1195
Db 1223 CTTCCACGCGGCTGATCAACGTCCACCCGACGCAAGTCAATCCACAGGCTTCTTTT 1282
QY 1196 AATCTGGCCAGTCTGAGAGCTCAGTAACTAGTAGTATAG 1240
Db 1283 AATCTGGCCAGTCTGAGAGCTCAGTAACTAGTAGTATAG 1327

RESULT 2

US-09-918-995-3342
; Sequence 3342, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3342
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(474)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-3342

Query Match 21.6%; Score 268; DB 10; Length 474;
Best Local Similarity 100.0%; Pred. No. 3.5e-129;
Matches 268; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 973 GAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGAATACTGAAACTC 1032
Db 63 GAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGAATACTGAAACTC 122
QY 1033 CGGCTACAAATTCAGATGAACAAAGACCAAGAAACCGCTTCTCTCCCTCCCAATCCTG 1092
Db 123 CGGCTACAAATTCAGATGAACAAAGACCAAGAAACCGCTTCTCTCCCTCCCAATCCTG 182
QY 1093 TGAATCCAGTGTCTCTGAAACAAAGAAACAGTAGCTGTGCGCTCTCAGCGGGCTGA 1152
Db 183 TGAATCCAGTGTCTCTGAAACAAAGAAACAGTAGCTGTGCGCTCTCAGCGGGCTGA 242
QY 1153 TCAACGCTCCACCGCCAGCAGCAAGTCATCCACAGGCTTCTTTTAATCTGCCAGTCTCTGA 1212
Db 243 TCAACGCTCCACCGCCAGCAGCAAGTCATCCACAGGCTTCTTTTAATCTGCCAGTCTCTGA 302
QY 1213 GAAGTCTAGTAAATACAATCTAGTATAG 1240
Db 303 GAAGTCTAGTAAATACAATCTAGTATAG 330

RESULT 3

US-09-864-761-16305
; Sequence 16305, Application US/09864761
; Patent No. US2002048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30

```
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16305
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF1211782.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
; US-09-864-761-16305

Query Match      11.6%; Score 144; DB 9; Length 401;
Best Local Similarity 100.0%; Pred. No. 2.4e-64;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 AGTCCTGAGGGCTCCAGGCTCGCTCACTGACCGTCTCCAGGCTGGAGGCTCAT 222
DB 258 AGTCCTGAGGGCTCCAGGCTCGCTCACTGACCGTCTCCAGGCTGGAGGCTCAT 317
QY 223 CATGTGGGCTCTCAGTGACATGTTGGTGAAGCGTCAGGCCCATGGAGCCCATCATCAC 282
DB 318 CATGTGGGCTCTCAGTGACATGTTGGTGAAGCGTCAGGCCCATGGAGCCCATCATCAC 377
QY 283 CAATGACCGCTTCACCTCTCAGAG 306
DB 378 CAATGACCGCTTCACCTCTCAGAG 401

RESULT 4
US-09-864-761-16653
; Sequence 16653, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16653
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF064857.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.93
; US-09-864-761-16653

Query Match      2.0%; Score 25; DB 9; Length 357;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 949 AGGATTTTCGTATTCAATTTCAAAAG 973
DB 42 AGGATTTTCGTATTCAATTTCAAAAG 66

RESULT 5
US-10-027-632-255100
; Sequence 255100, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 255100
; LENGTH: 1078
; TYPE: DNA
```

```
; ORGANISM: Human
US-10-027-632-255100

Query Match
Best Local Similarity 1.9%; Score 24; DB 13; Length 1078;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AGTTTCCTAGGCTGCCATAACAAA 53
    |||||
Db 871 AGTTTCCTAGGCTGCCATAACAAA 894

RESULT 6
US-10-027-632-255101
; Sequence 255101, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 255101
; LENGTH: 1078
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-255101

Query Match
Best Local Similarity 1.9%; Score 24; DB 13; Length 1078;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AGTTTCCTAGGCTGCCATAACAAA 53
    |||||
Db 871 AGTTTCCTAGGCTGCCATAACAAA 894

RESULT 7
US-10-027-632-255102
; Sequence 255102, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
```

```
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 255102
; LENGTH: 1078
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-255102

Query Match
Best Local Similarity 1.9%; Score 24; DB 13; Length 1078;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AGTTTCCTAGGCTGCCATAACAAA 53
    |||||
Db 871 AGTTTCCTAGGCTGCCATAACAAA 894

RESULT 8
US-10-027-632-255100
; Sequence 255100, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 255100
; LENGTH: 1078
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-255100

Query Match
Best Local Similarity 1.9%; Score 24; DB 16; Length 1078;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AGTTTCCTAGGCTGCCATAACAAA 53
    |||||
Db 871 AGTTTCCTAGGCTGCCATAACAAA 894

RESULT 9
US-10-027-632-255101
; Sequence 255101, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
```

FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 255101
; LENGTH: 1078
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-255101

Query Match 1.9%; Score 24; DB 16; Length 1078;
Best Local Similarity 100.0%; Pred. No. 0.14; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0;

QY 30 AGTTTCCTAGGCTGCCATACAAA 53
|||||
Db 871 AGTTTCCTAGGCTGCCATACAAA 894

RESULT 10

US-10-027-632-255102
; Sequence 255102, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 255102
; LENGTH: 1078
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-255102

Query Match 1.9%; Score 24; DB 16; Length 1078;
Best Local Similarity 100.0%; Pred. No. 0.14; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0;

QY 30 AGTTTCCTAGGCTGCCATACAAA 53
|||||

Db 871 AGTTTCCTAGGCTGCCATACAAA 894

RESULT 11

US-10-027-632-257899
; Sequence 257899, Application US/10027632
; Publication No. US20030198371A1
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 257899
; LENGTH: 1121
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)...(1121)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-257899

Query Match 1.9%; Score 24; DB 13; Length 1121;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AGTTTCCTAGGCTGCCATACAAA 53
|||||
Db 915 AGTTTCCTAGGCTGCCATACAAA 938

RESULT 12

US-10-027-632-257899
; Sequence 257899, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 257899
; LENGTH: 1121
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(1121)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-257899

Query Match 1.9%; Score 24; DB 16; Length 1121;
Best Local Similarity 100.0%; Pred. No. 0.14; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0;

QY 30 AGTTTCCTAGGTGCATACAAA 53
DB 915 AGTTTCCTAGGTGCATACAAA 938
|||||

RESULT 13
US-10-437-963-81852/c

; Sequence 81852, Application US/10437963
; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 81852

; LENGTH: 1284

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_81337C.1

US-10-437-963-81852

Query Match 1.9%; Score 23; DB 17; Length 1284;
Best Local Similarity 100.0%; Pred. No. 0.46; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0;

QY 872 CGCTGCTGCTGTCGCCGCTGC 894
DB 156 CGCTGCTGCTGTCGCCGCTGC 134
|||||

RESULT 14

US-09-770-445-930

; Sequence 930, Application US/09770445

; Patent No. US20020023281A1

; GENERAL INFORMATION:

; APPLICANT: Gorlach, Jorn

; APPLICANT: An, Yong-Qiang

; APPLICANT: Hamilton, Carol M.

; APPLICANT: Price, Jennifer L.

; APPLICANT: Raines, Tracy M.

; APPLICANT: Yu, Yang

; APPLICANT: Rameaka, Joshua G.

; APPLICANT: Page, Amy

; APPLICANT: Matthew, Abraham V.

; APPLICANT: Ledford, Brooke L.

; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kriker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis

; FILE REFERENCE: thaliana

; CURRENT APPLICATION NUMBER: US/09/770,445

; CURRENT FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: US 60/178,472

; PRIOR FILING DATE: 2000-01-27

; NUMBER OF SEQ ID NOS: 999

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 930

; LENGTH: 767

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-770-445-930

Query Match 1.8%; Score 22; DB 9; Length 767;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 886 CCGCGCTGCTGTTGTTGGCTGC 907
DB 672 CCGCGCTGCTGTTGTTGGCTGC 693
|||||

RESULT 15

US-09-220-920-65

; Sequence 65, Application US/09220920

; Patent No. US2002002269A1

; GENERAL INFORMATION:

; APPLICANT: Milbrandt, Jeffrey D.

; APPLICANT: Baloh, Robert H.

; TITLE OF INVENTION: Artemin, A No. US20020002269A1el Neurotrophic Factor

; FILE REFERENCE: 6029-7996

; CURRENT APPLICATION NUMBER: US/09/220,920

; CURRENT FILING DATE: 1998-12-24

; EARLIER APPLICATION NUMBER: 09/163,283

; EARLIER FILING DATE: 1998-09-29

; EARLIER APPLICATION NUMBER: 60/108,148

; EARLIER FILING DATE: 1998-11-12

; EARLIER APPLICATION NUMBER: 09/218,698

; NUMBER OF SEQ ID NOS: 120

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 65

; LENGTH: 1203

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-220-920-65

Query Match 1.8%; Score 22; DB 9; Length 1203;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 873 GCTGCTGCTGCTGCCGCTGC 894
DB 51 GCTGCTGCTGCTGCCGCTGC 72
|||||

Search completed: September 15, 2004, 07:20:17
Job time : 659.934 secs